

Tilburg University

Value-Addes Tax and Shadow Economy

Hoseini, Mohammad

Publication date:
2015

[Link to publication in Tilburg University Research Portal](#)

Citation for published version (APA):

Hoseini, M. (2015). *Value-Addes Tax and Shadow Economy: the Role of Input-Output Linkages (revision of CentER Discussion Paper 2013-036)*. (CentER Discussion Paper; Vol. 2015-058). Economics.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

No. 2015-058

**VALUE-ADDED TAX AND SHADOW ECONOMY:
THE ROLE OF INPUT-OUTPUT LINKAGES**

By

Mohammad Hoseini

14 December 2015

This is a revised version of CentER Discussion Paper
No. 2013-036

ISSN 0924-7815
ISSN 2213-9532

Value-Added Tax and Shadow economy: the Role of Input-Output Linkages

Mohammad Hoseini*

December 11, 2015

Abstract

Under the VAT, formal traders report their purchases to the administration for a deduction in their VAT bill. This paper models this third-party reporting feature of the VAT in an input-output economy and quantifies it among different activities using a forward linkages index. The administration can reduce the size of shadow economy by reallocating visiting audits to backwardly linked activities and cross-checking VAT payments with input credit claims in forwardly linked activities. Empirical evidence from Indian service sector justifies the assumptions and suggests a significant increase in the tax compliance of forwardly linked activities following VAT adoption in 2003.

Keywords: Value-added tax, Informality, Tax enforcement, Linkage analysis
JEL code: H26

1 Introduction

Over the past decades, more than 150 countries, including many of the poorest with large shadow economies,¹ has adopted the value-added tax (henceforth VAT). Shadow economy hinders the efficiency of the VAT (Piggott and Whalley, 2001; Emran and Stiglitz, 2005; Keen, 2008), but on the other hand the third-party reporting property of the VAT makes evasion and informality harder (Kopczuk and Slemrod, 2006; Pomeranz, 2015). Under the input refund system of the VAT, formal enterprises desire to receive VAT credit on inputs and, as a result, inform the tax administration about their purchases. This incentive, also labeled as “self-enforcing” mechanism, provides third-party reported records on sales to formal clients and helps the administration to find informal suppliers. But, if all clients are informal firms

*Tilburg University, PO Box 90153 5000 LE Tilburg, The Netherlands; Sharif University, Tehran, Iran, M.Hoseini@uvt.nl. I am grateful to Jan Boone and Jenny Ligthart[†] for their helpful suggestions. I have also benefited from comments by Thorsten Beck and seminar participants at Tilburg University.

¹Up to 2012, 26 low income, 40 lower middle income, 46 upper middle income, and 38 high income countries implemented the VAT (Keen, 2013). According to Schneider et al. (2011), on average the size of shadow economy in developing countries is around 40% and ranges up to 70%.

or final consumers, no record of purchase is reported and the mechanism breaks down. Hence, the VAT changes in third-party reporting across different activities based on the formality status of the clients.

To flash the variation in VAT self-enforcement across activities, I present a preliminary evidence from India.² The government of India adopted VAT on services in 2003. Figure 1 illustrates tax registration rate among firms in different services versus the share of final consumption in each activity. The \circ and \times symbols respectively represent the percentage of enterprises registered for tax in 2001 and 2006. A negative relationship between tax registration rate and the share of final consumption in an activity is seen only after and not before VAT adoption. In other words, activities with more sales to businesses rather than households experience a higher tax compliance when the VAT is in place.

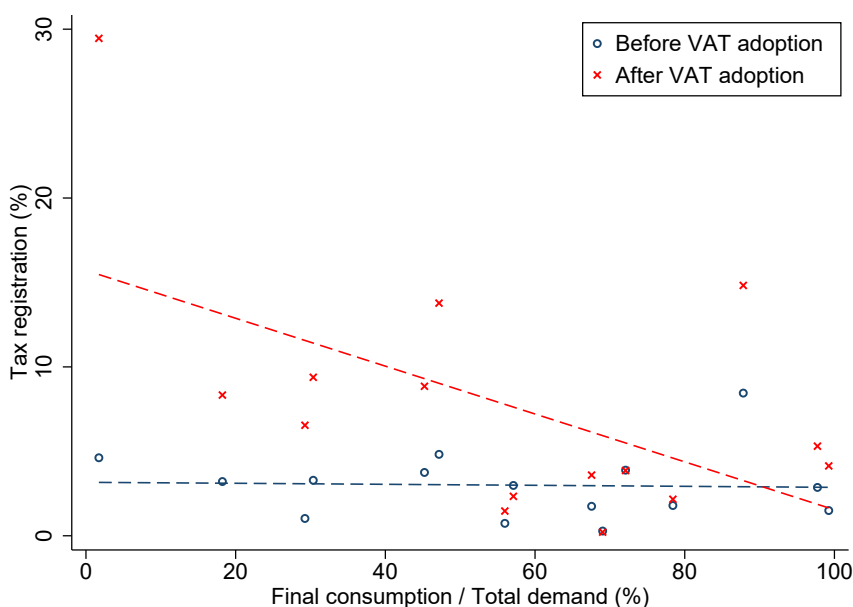


Figure 1: Tax registration rate and final consumption in different service sectors of India before and after VAT adoption in 2003. The \circ and \times symbols represent the registration rates in 2001 and 2006 respectively. The estimated slope for 2006 is -0.14 with standard deviation 0.059. The registration rate is measured for a representative sample of “unorganized service sector”. See section 5 for definitions and data details.

Despite the prolonged recognition of the third-party reporting feature of the VAT (e.g. Tait, 1972, 1991), it has only recently been empirically tested by Pomeranz (2015). She reports that a random audit letter, in addition to increasing the VAT payment of the treated firm, also increases tax compliance of the suppliers but not the clients. Yet, the variation in VAT self-enforcement across activities and an index for measuring it is absent from the

²In section 5, I describe the data and the reform and test this in a more robust fashion.

literature. This paper tries to tackle two questions in this regard: how does VAT affect tax evasion across different activities? and how should VAT audit policies be designed across different activities in an economy characterized by a large informal sector? My approach to answer these questions is based on exogenous variation in forward and backward linkages between activities.

Although the role of client's formality status on VAT evasion is intuitive, determining the degree of third-party reporting in each activity and gauging the spill-overs on their suppliers are not straightforward. In reality, the economy is composed of very sophisticated input-output structures with complex linkages. Some goods are usually sold only to firms or only to households (e.g. mining activities vs. child caring services), but a wide range of goods are purchased by both firms and households (financial services, transportation, information technology, etc.). Moreover, even if the enterprise nature of the customers is known, to measure VAT self-enforcement, formal and informal business clients must be distinguished. The position of the activity in the production chain and linkages with other activities are the main factors to determine the type of clients and the spillovers of the transmission of formality to the suppliers.

For this reason, I build an input-output model which lays the groundwork for defining *forward* and *backward* linkages. These two sectoral indices respectively reflect the flow of intermediate products *to* and *from* other activities. Next, I add the VAT to this structure and explain visiting audits and cross-checking invoices – matching each input credit claim with a corresponding output tax payment – as two distinct enforcement tools in the VAT system. The theoretical results suggest that in backwardly linked activities, which buy their inputs from other activities, tax administrators should focus on revealing internal information of firms by visiting audits. In comparison, in the forwardly linked industries the administration should concentrate on arm's-length transactions and cross-checking the invoices with the corresponding input credit claims.

Using firm-level service sector data from India in 2001 and 2006, the implications of the model are examined by taking the VAT adoption in 2003 as a natural experiment and comparing formality in each activity before and after it. Formality is estimated based on tax registration rate and share of production by taxpayer firms in each activity. I also compute forward and backward linkages of each activity using the input-output tables of India in 2003. The results confirm a positive relationship between activity's forward linkages and formality

after VAT adoption. For example, after VAT adoption in 2003, registration rate of storage and warehousing activities (forward linkages=2.78) increased by 8.97 percentage point more than of training and coaching (forward linkages=1.01).

The theoretical contribution of this paper is formalizing the self-enforcing feature of the VAT. For a complex network of activities, the paper provides a reduced-form formula for VAT enforcement policy by accounting for not only the revenue gains of less evasion in each activity but also the spill-over gains of the compliance of its suppliers. The paper also fills a gap in the tax evasion theory by providing a tractable model of intersectoral linkages to study the tax enforcement problem. Despite the long-established literature on linkage analysis (which dates back to Rasmussen, 1957), its applications in the economic literature is, for the most part, neglected by theorists. The empirical contribution of this paper is providing a novel gauge for each activity to formulate VAT enforcement policy. The great advantage of the linkages indices is the simple measurability using input-output tables easing policy implementation in developing countries. Testing the application of the provided indices in India confirms their validity to quantify the level of formality transmission and formulate the enforcement policy.

In this paper, informality refers to non-registration under tax administration. There are a number of other frauds in the VAT which are discussed in details in Keen and Smith (2006). In high-income countries, underground activities comprise relatively smaller share in total VAT evasion rather than the frauds by registered traders such as misreporting the transactions, carousel fraud, etc.³ In comparison, shadow economy fraud or failure to register is common in low-income countries. For instance, in section 5 I find that during 2001-2006 in India, only 17 percent of production in “unorganized service sectors” is produced by tax registered firms. Hence, the results of this paper is more applicable for developing countries suffering from large shadow economies.

The findings of the paper contribute to the public finance literature in several ways. A number of studies in public economics model the VAT from different angles. Some, on a broader level, compare the efficiency of the VAT to other indirect taxes like tariff and RST (e.g. Emran and Stiglitz, 2005; Keen, 2008; Boadway and Sato, 2009), while others concentrate on its optimal design of rates and exemption levels (e.g. Piggott and Whalley, 2001; Keen and Mintz, 2004). To my knowledge, the only research presenting a theory of VAT evasion is

³As an example, in 2001-02, VAT revenue loss in UK is estimated to be %10.2–%14.2 of total VAT revenue, among which %0.6–%0.7 is due to non-registration for the VAT (Keen and Smith, 2006). The misreporting fraud – understating sales and overstating purchases – by registered firms is studied in Hoseini (2014) in details.

De Paula and Scheinkman (2010) who address the decision of firms for being formal in a VAT setting. They consider a two stage production chain with firms heterogeneous in terms of managerial ability and show that (in)formal firms trade with their (in)formal peers in the production chain not with the other type. In my model, the same assortative equilibrium can hold, but I show that an alternative equilibrium with trade between the two sectors is also possible. The model of De Paula and Scheinkman (2010) is silent about the activity of the firm and the linkages.

This paper also contributes to the growing literature that studies the role of firms in facilitating tax enforcement. Kopczuk and Slemrod (2006) emphasize the importance of inter-firm information and argue that verifiable paper trails from arm's-length transactions and aggregation of information within firms considerably improves tax enforcement. Kleven et al. (2015), on the other hand, focus on intra-firm information and highlight the role of firms in facilitating tax enforcement by providing information through third-party reporting of employees. The VAT is paid by firms and thus enables the tax administration to use both inter-firm and intra-firm information sources for detecting evasion. To formulate detection probability, this paper puts together these two instruments and assume one part of detection is made by checking arm's-length transactions and the other by third-party reporting of workers. Finally, this paper contribute to the literature about taxation in developing countries (e.g. Gordon and Li, 2009; Best et al., 2016). While most part of this literature addresses optimal taxation in the presence of limited tax capacity, my paper derives a reduced-form formula to increase tax capacity in a developing country.

This paper is organized as follows. Section 2, constructs the basic model and introduces forward and backward linkages. Next, in section 3, I explain tax administration's enforcement instruments in the VAT, and in section 4, I find the market equilibrium and the optimal enforcement policy in each industry. In the section 5, I empirically examine the existence of the self-enforcing feature and finally, I conclude.

2 The Basic Model

Consider an economy with $n = 1, \dots, N$ competitive industries each producing a homogeneous and unique product. Each product is produced using intermediate and non-intermediate inputs and can be used as an input of other industries or be consumed by final consumers.

Given this general picture, I assume an analytically tractable structure that can be connected to the input-output tables of an economy. I first describe the economic environment and then characterize the competitive equilibrium.

The production technology of all industries is a Cobb-Douglas function as follows⁴

$$x_n = v_n^{\alpha_n} \prod_{k=1}^N x_{kn}^{\alpha_{kn}}, \quad \sum_{k=1}^N \alpha_{kn} + \alpha_n = 1 \quad (1)$$

where x_n is the production of industry n , v_n is the value-adding input,⁵ and x_{kn} is the amount of product k as an intermediate input of industry n . The factor coefficients α_n and α_{kn} determine the technology in each industry and correspond to input-output coefficients. Higher $\sum_k \alpha_{kn}$ reflects that the industry n relies more on intermediate inputs than value-adding inputs. In comparison, higher $\sum_k \alpha_{nk}$ reflects more consumption of product n as intermediate input than final good.

The market of the value-adding input v is competitive and I take its price as the numeraire. In equilibrium, p_n is the price by which the product n is traded. Under perfect competition, the profit of a representative firm in industry n becomes

$$\pi_n = p_n x_n - \sum_{k=1}^N p_k x_{kn} - v_n \quad (2)$$

and the first order conditions result

$$x_{kn} = \frac{\alpha_{kn} p_n}{p_k} x_n, \quad v_n = \alpha_n p_n x_n \quad (3)$$

By substituting (3) in the production function (1), we can find the relationship between prices as

$$p_n = \left(\frac{1}{\alpha_n}\right)^{\alpha_n} \prod_{m=1}^N \left(\frac{p_m}{\alpha_{mn}}\right)^{\alpha_{mn}} \quad (4)$$

At the demand side, each product is used either as an intermediate input or for final consumption. By denoting the final consumption of the product of industry n as x_{0n} , one can

⁴The results of the model is robust to assuming a general CES function.

⁵For simplicity, v_n represents all inputs other than intermediate goods. Because, my final goal is modeling VAT, intermediate goods correspond to taxable goods and services. More general assumption would be adding labor and other non-intermediate factors of value-added in the production function and using $l_n^{a_1} k_n^{a_2} \dots$ instead of v_n , but it complicates the model without changing the qualitative results.

write total representative demand of industry n as

$$x_n = x_{0n} + \sum_{k=1}^N x_{nk} \quad (5)$$

where the first term shows the final demand and the second term is the sum of all intermediate demands of product n . Assume that final consumers are homogeneous with a net income w and a Cobb-Douglas utility in the sense that the utility maximization problem for a representative final consumer is

$$\max u(x_{01}, \dots, x_{0N}) = \sum_{i=1}^N \beta_i \ln(x_{0i}), \quad \sum_{i=1}^N \beta_i = 1, \quad \text{s.t.} \quad \sum_{i=1}^N p_i x_{0i} = w \quad (6)$$

where β_i determines per-unit fraction of the consumers' incomes used in purchasing good i . The solution of (6) gives the final demand of each product n as $x_{0n} = w\beta_n/p_n$.

The representative production and demand are equal in all industries at equilibrium and by substituting x_{0n} and x_{nk} from (3) in (5), we can write

$$x_n = w\beta_n/p_n + \sum_{k=1}^N \frac{\alpha_{nk}p_k}{p_n} x_k, \quad \Rightarrow \quad p_n x_n = w\beta_n + \sum_{k=1}^N \alpha_{nk} p_k x_k \quad (7)$$

Define $y_n = p_n x_n$ as representative value of production in industry n . Then,

$$y_n = w\beta_n + \sum_{k=1}^N \alpha_{nk} y_k \quad (8)$$

(8) expresses that the total value of production of commodity n is equal to the value that is bought by final consumers, who spend a fraction β_n of their total wealth on it, plus the intermediate demand of other industries. The share of the value of product n consumed as the input of k in total value of output of k is α_{nk} . To find the equilibrium value of production in each industry, I employ vectors Y and B with dimensionality $N \times 1$ containing y_n and β_n respectively, together with an $N \times N$ matrix \mathbf{A} built by α_{kn} as elements.⁶ Then, (8) is written as

$$Y = wB + \mathbf{A}Y, \quad \Rightarrow \quad Y = (\mathbf{I} - \mathbf{A})^{-1}wB. \quad (9)$$

Since $wB > 0$, (9) imposes that $\mathbf{A}Y < Y$. In this case, $(\mathbf{I} - \mathbf{A})^{-1}$ in (9) always exists meaning

⁶Throughout the paper, lower-case letters like x denotes a scalar, upper-case like X a vector, and bold upper-case like \mathbf{X} a matrix.

there is unique equilibrium Y and we have⁷

$$(\mathbf{I} - \mathbf{A})^{-1} = \sum_{m=0}^{\infty} \mathbf{A}^m$$

Therefore, $Y = (\mathbf{I} + \mathbf{A} + \mathbf{A}^2 + \dots)wB$, where $\mathbf{A}^m wB$ is the indirect effect of final consumption wB in rising the value of production of Y through m intermediate goods in between. In the following, I first introduce the concepts of forward and backward linkages, then in the next section, I model formal and informal sector in this economy.

2.1 Forward and backward linkages

The above model provides a simple framework to study the effect of inter-sectoral linkages. I introduce two types of linkages indices in this regard: *forward linkage* which quantifies the value of products that is *sold to* the other industries (not final consumers), and *backward linkage*, which indexes the value of products *purchased from* other industries. Figure 2 shows a simple example of a network of 6 industries, where there are two vertical chains, each including a raw material supplier (S), a manufacturer (M), and a retailer (R), such that S is the upstream and R is the downstream industry. In other words, in each chain S is just forwardly linked, R just backwardly and M has both types of linkages. These two production chains are also horizontally linked from the middle. Therefore, the upstream industry S_1 ,

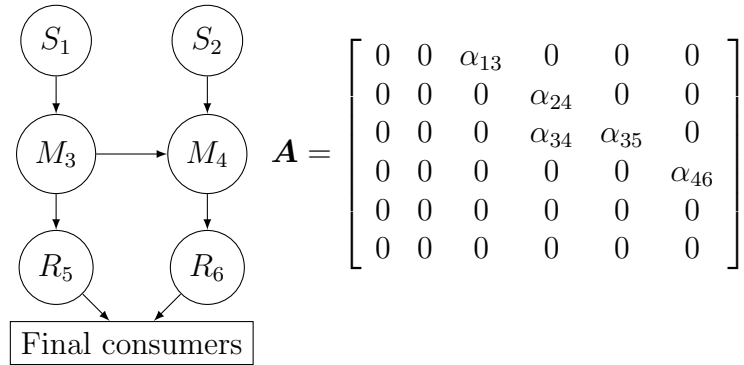


Figure 2: A simple industry network and the corresponding input-output table \mathbf{A} . Industry S_1 is forwardly linked to M_3 (direct linkage) and M_4 , R_5 , and R_6 (indirect linkages). In comparison, Industry R_6 is backwardly linked to M_4 (direct linkage) and M_3 , S_1 , and S_2 (indirect linkages).

in addition to M_3 and R_5 , is forwardly linked with M_4 and R_6 through M_3 . Similarly, the backward linkages of R_6 include the direct linkage of M_4 and indirect linkages of S_1 , S_2 and M_3 through M_4 .

⁷See Ten Raa (2006), Theorem 2.2.

We can gauge these two properties using input-output coefficients. For simplicity assume that the production value in all industries is one. Then, direct forward linkage of n to k – which is also the direct backward linkages of k to n – is reflected in α_{nk} . Similarly, the indirect linkages from n to k through industry j can be gauged by $\alpha_{nj}\alpha_{jk}$. Therefore in Figure 2, a measure for the forward linkages of S_1 would be $\alpha_{13} + \alpha_{13}\alpha_{34} + \alpha_{13}\alpha_{35} + \alpha_{13}\alpha_{34}\alpha_{46}$, which is the increase in production of S_1 in response to one unit increase in the production of all other industries. Similarly the backward linkages of R_6 can be gauged by $\alpha_{46} + \alpha_{24}\alpha_{46} + \alpha_{34}\alpha_{46} + \alpha_{13}\alpha_{34}\alpha_{46}$ as the increase in the production of all other industries in response to one unit increase in the production of R_6 .

A general formula for forward and backward linkages is first introduced by Rasmussen (1957). He suggests using the row and column sums of the Leontief inverse of the industry \times industry table of the economy which is computed using the corresponding input-output transaction tables. Using the same approach, I define the following formulas for forward and backward linkages vectors:

$$\text{Forward linkage: } FL = (\mathbf{I} - \mathbf{Y}^{-1}\mathbf{A}\mathbf{Y})^{-1}J \quad (10)$$

$$\text{Backward linkage: } BL = (\mathbf{I} - \mathbf{A}')^{-1}J \quad (11)$$

where \mathbf{Y} is a diagonal matrix of vector Y in the sense that $\text{diag}(\mathbf{Y}) = Y$, \mathbf{A}' is the transpose of \mathbf{A} , and J is a vector of ones (summation vector). If we expand (10) and (11) for each element, it gives the coefficients of each industry n as

$$f_l_n = 1 + \underbrace{\sum_{k=1}^N \alpha_{nk} \frac{y_k}{y_n}}_{\text{direct forward linkages}} + \overbrace{\sum_{k=1}^N \sum_{j=1}^N \alpha_{nk} \alpha_{kj} \frac{y_j}{y_n}}^{\text{indirect forward linkages through one intermediate good}} + \dots \quad (12)$$

$$bl_n = 1 + \underbrace{\sum_{k=1}^N \alpha_{kn}}_{\text{direct backward linkages}} + \overbrace{\sum_{k=1}^N \sum_{j=1}^N \alpha_{jk} \alpha_{kn}}^{\text{indirect backward linkages through one intermediate good}} + \dots \quad (13)$$

f_l_n shows how much is the value of production that goes to the other industries and not final consumers (the sum of all $\alpha_{nk}y_k, \alpha_{nk}\alpha_{kj}y_j, \dots$), normalized by y_n . It includes both direct linkages (customers) and indirect linkages through intermediate goods (customers of

customers). Similarly bl_n measures the value of inputs that comes from other industries, directly or indirectly (the sum of all $\alpha_{kn}y_n, \alpha_{jk}\alpha_{kn}y_n, \dots$), normalized by y_n . The important difference between fl_n and bl_n is that “ n ” is the first subscript in all terms of fl_n reflecting the output flow to other industries, while in bl_n , “ n ” is the last subscript reflecting the input requirements from other industries.

Note that these two indices are independent of each other. There may be an industry that has strong linkages in both directions or none of them. For instance, some primitive agricultural products or household activities are directly consumed to the final consumers. Hence, they possess no forward or backward linkages. In comparison, steel is a good example of an industry that has both types of linkages: coal and iron ore are backward and many items like canned goods are its forward industries. Furthermore, the two above linkage concepts cover both vertical and horizontal inter-industry connections. The vertical linkage is a chain of industries from up to downstream producing a final good (e.g. petroleum industries from crude oil extraction to plastics), but horizontal linkage is between two industries exchanging their differentiated products and none of them is necessarily categorized above or below the other (e.g. Fuels and Iron industries). The great advantage of the above indices is the inclusion of all types of linkages of a complex industry network in two simple and measurable formulas (10) and (11).

3 Adding the VAT and modeling informality

Now return to the basic model and consider the government imposes a uniform VAT rate t in all industries and fines all non-registered firms that are detected. Formal (informal) status is then defined as registered and fully tax compliant (not registered) for the VAT.⁸ In order to characterize formal and informal sectors, I follow the framework used by Piggott and Whalley (2001), Keen (2008), and Stiglitz (2009). In each industry, consider two formal and informal markets that both are competitive and produce the same product. Formal firms have constant returns and informal firms have decreasing returns due to the expected fine for VAT evasion. The simplifying assumption used in this structure is the absence of occupational

⁸Although this definition ignores the possibility of misreporting in the formal sector, one can assume each representative firm sells one part of its production in the formal market (x_n^f) and one part in the informal (x_n^i). In this case, as shown below, if the expected fine of producing x_n^i is the same as an informal firm with that level of production, the results do not change. For detailed analysis of misreporting in the VAT see Hoseini (2014).

choice and transition between the two sectors. In fact, the economy is considered to be composed of two types of producers with high and low abilities who can only work as formal and as informal respectively. In this setting, the two perfectly competitive sectors can co-exist in all industries and response in size to the tax enforcement policy. The advantage of this assumption is generating analytically tractable results in my model.⁹

Under the VAT, the ad valorem tax in industry n is added to its market price p_n , in the sense that formal firms charge tax on their sales and issue the corresponding invoices for the buyers, who if registered can use this invoice to refund against their own tax liability. Informal firms, on the other hand, buy and sell at tax inclusive price, but they do not pay any tax to the government.

As an example, consider a simple production chain for a final good R with three stages: raw material supplier (S), manufacturer (M), and retailer (R) such that their pre-tax prices are p_S , p_M , and p_R respectively. We can illustrate the value chain as $0 \xrightarrow{S} p_S \xrightarrow{M} p_M \xrightarrow{R} p_R$. Therefore, at the first stage, one unit of S transforms to one unit of M and then, at the second stage, the new product changes to one unit of final consumer good R. Table 1 shows the balance sheet of each industry if they register under the VAT or not.

Table 1: Prices and revenues under the VAT.				
industry	output price	input price	tax payment	profit
S^f	$p_S + tp_S$	0	$tp_S x_S^f$	$p_S x_S^f$
M^f	$p_M + tp_M$	$p_S + tp_S$	$t(p_M x_M^f - p_S x_S^f)$	$p_M x_M^f - p_S x_S^f$
R^f	$p_R + tp_R$	$p_M + tp_M$	$t(p_R x_R^f - p_M x_M^f)$	$p_R x_R^f - p_M x_M^f$
S^i	$p_S + tp_S$	0	0	$(1+t)p_S x_S^i - c_S$
M^i	$p_M + tp_M$	$p_S + tp_S$	0	$(1+t)(p_M x_M^i - p_S x_S^i) - c_M$
R^i	$p_R + tp_R$	$p_M + tp_M$	0	$(1+t)(p_R x_R^i - p_M x_M^i) - c_R$

The superscript f represents formal sector and i informal. The difference between the two groups is that informal firms do not pay the tax to the government, but instead have a risk of detection and punishment by the tax administration equal to c_n . In the general model with N industries the same pattern holds. According to Table 1, tax earnings cancel out the

⁹Another way of modeling informality is assuming heterogeneous producers with different abilities who can freely choose to be formal or informal (Rauch, 1991; Fortin et al., 1997; De Paula and Scheinkman, 2010). In these models, there is a size constraint for informal firms that leads to a cut-off ability above which producers become formal and vice-versa. For illustrative purposes I do not choose this approach, because as I discuss below, it complicates the analysis without providing any additional insight.

payments of a formal firm to the government, therefore its profit is tax-exclusive. Then,

$$\pi_n^f = p_n x_n^f - \sum_{k=1}^N p_k x_{kn}^f - v_n^f, \quad x_n^f \geq 0 \quad (14)$$

which is similar to (2) and gives (4) as equilibrium prices.

Table 1 shows that informal firms benefit from evading the VAT, but bear a cost c_n . This cost reflects the risk of being detected and punished by the tax administration due to not registering for the VAT. I assume the evasion cost is a convex function of the level of production x_n^i , and thus the production of informal firms is subject to decreasing returns reflecting the expected loss from VAT evasion.¹⁰ Following Allingham and Sandmo (1972), the literature on tax evasion normally assumes that the cost of evasion c_n depends on two factors. One is the value of the fine, reasonably proportional to the evasion level, and the other is the probability of detection the fraud (for a survey see Slemrod and Yitzhaki, 2002). In this model, I include these properties in the following simple form:

$$c_n(e_n, \eta_n) = \theta \eta_n e_n, \quad 0 \leq \eta_n \leq 1 \quad (15)$$

where η_n and e_n are detection probability and evasion level, respectively. $\theta > 1$ is the punishment per unit of evasion set by the government based on legal and political considerations. It should be greater than one, otherwise the marginal benefit of informality always beats its costs. Therefore, the profit function of a representative firm in the informal sector of industry n is written as

$$\pi_n^i = (1+t)p_n x_n^i - \sum_{k=1}^N (1+t)p_k x_{kn}^i - v_n^i - \theta \eta_n e_n, \quad x_n^i \geq 0. \quad (16)$$

Proposition 1 *Under perfect competition, the production of a representative firm in the in-*

¹⁰As mentioned above, this assumption allows for the coexistence of formal and informal markets in the sense that the informal sector produce at a specific level and the rest of demand is supplied by formal firms. Boadway and Sato (2009) and De Paula and Scheinkman (2010) endogenize the occupational choice by assuming that formal status has positive impact on firm's profit due to access to formal services (e.g. financial intermediation, access to capital market, hiring skilled workers), and arguing the extent of these advantages depends on the managerial ability of the firm. In this model, one can consider a same structure with a continuum of producers with different abilities and assume that the ability adds a new term to the profit in the formal sector. Then above a cut-off ability, firms become formal and vice versa. This assumption however is not determinant and do not alter the theoretical predictions of the model.

formal sector of industry n is obtained from

$$x_n^i \frac{\partial \eta_n}{\partial x_n^i} + \eta_n = \frac{1}{\theta} \quad (17)$$

Proof 1 *Appendix A.1.*

According to Proposition 1, to compute the informal production of each industry, we need to know the functional form of detection probability η_n . In the next section, I discuss in detail how we can mathematically define this function and link it with the self-enforcing feature of the VAT. After finding the representative production of the informal sector from Proposition 1, we have $y_n^i = p_n x_n^i$. Then, the representative formal production in each industry can be computed as $y_n^f = y_n - y_n^i$, where the representative total production y_n is determined by input-output equation (9). The desired formality index for the tax administration is y_n^f/y_n which is the share of formal production in each industry.

3.1 Probability of detection

Before any further analysis, we need a reasonable assumption about the probability of detection η_n . Broadly, there are two information sources for the tax administration to check the business records. One is within-firm information such as profit and wages, the other is external information like transactions with other businesses. Detection of within-firm information is usually done by visiting audit, but the external information is attainable by cross-checking the transaction records of business partners. In the following, I first explain these factors and how the administration can exploit them and then I define a functional form for this probability.

3.1.1 Firm size

A crucial factor that determines tax evasion is the firm size. Kleven et al. (2015) show that although theoretically a firm can collude with its employees to unreport its tax liability, in practice, if the number of employees is large, maintaining such collusion becomes very difficult. They indicate this collusion may break for a number of reasons. The first set is a random shock due to unplanned events such as a conflict between employer and employee, moral concerns of a newly hired employee, or a mistake of an employee or employer to reveal the records to the tax administration. The other reason can be rational whistle blowing because

of a rewards by the tax administration which is common in several countries.¹¹ In addition, as an enterprise grows, the possibility of evasion decreases. Therefore, this type of third-party reporting dramatically improves tax enforcement of bigger firms. To incorporate this effect in the probability function of an informal firm, I use the size of the firm x_n^i . To ensure that the probability always lies between zero and one, I divide this term by \bar{x} which represents the size in which the firm is always detected. Hence, x_n^i/\bar{x} reflects the chance of detection by a visiting audit.

3.1.2 Firm-to-firm transactions

The second source of information for the tax administration is arm's-length transactions of the firm (Kopczuk and Slemrod, 2006). This source is especially important in the VAT because formal firms are required to keep the records of their sales and purchases, enabling the administration to verify them with the accounts of their business partners. In practice, there are different methods of VAT administration based on the type of assessment, technology and invoice reporting, but in all types, firms either have to keep the sales and purchase invoices for some years or send them to the tax administration with their tax report.¹² Regardless of the methods of administration, formal traders always have a strong incentive to report their purchase invoices, because they can deduct input taxes from their VAT bill. Each purchase invoice, as a third-party reported piece of information, enables the authority to investigate formality and registration status of the seller. In other words, the incentive of formal clients to report the transaction for getting input credit makes an additional risk of detection for the informal supplier.

This feature of the VAT, which is labeled as self-enforcing property (Keen and Smith, 2006), is a one-way effect from formal firms to their supplier not costumers. In her “spillover experiment”, Pomeranz (2015) reports that when the Chilean tax authority randomly sends an announcement of audit to half of the suspected firms, the announcement reduces the evasion among the suppliers but not clients of the treated firms. The reason is that when a formal firm sells a product, it charges VAT and issues the invoice irrespective of the buyer's type. If the buyer is a formal firm, it uses the invoice for input credit and the two records – tax

¹¹Kleven et al. (2015) mention some example from OECD countries like the US and Japan.

¹²With the new developments in electronic invoicing, in many countries firms report each single invoice to the tax administration electronically with substantially less compliance cost. This enables the administration to perform cross-checking without visiting the firm. For more information on different types of VAT administration see Ebrill et al. (2001) and Hoseini (2014).

payment and formal buyer's input credit claim – enables the tax administration to cross-check them together. Otherwise, the administration only receives a payment and be informed that the product is sold, regardless of whether the buyer is an informal firm or a final consumer. Thus, when the seller is formal but the buyer is not, the tax administration cannot distinguish between the ones that should be punished (informal firms) and not punished (final consumers).

To incorporate the self-enforcing feature in the detection probability, I assume that the risk of being detected by invoice cross-checking is positively associated with the share of formal customers, who inform the tax administration about the transaction. Hence, in addition to firms size, the detection probability is an increasing function of the share of sales to formal firms in total sales of industry n which is $\sum_{k=1}^N x_{nk}^f / x_n$. Notice that because both markets are perfectly competitive, a single informal firm cannot change x_n by changing its own production level and takes it as given when optimizing profit.

To formulate detection probability η_n , we need to specify the policy variables and define a function to put size and cross-checking risks together. The tax administration has two types of policies to make: 1) the allocation of enforcement expenditure across industries, 2) the allocation of enforcement type (vising audit versus cross-checking) within each industry. I assume that the administration's policy in each industry is primarily reflected in parameter $\phi_n = g_n / \bar{g}$, where g_n is the expenditure to enforce tax in industry n and \bar{g} is a fixed term to ensure $0 \leq \phi_n \leq 1$. For the first policy, the administration seeks to find the optimal allocation of g_n – and consequently ϕ_n – across different industries for maximizing net VAT revenue. The second policy parameter is the share λ_n of the enforcement ϕ_n to dedicate for invoice cross-checking which means how much the administration tries to find and check the seller of a purchase invoice claimed by a formal customer. Therefore, $\lambda_n \phi_n$ determines the level of invoice cross-checking and $(1 - \lambda_n) \phi_n$ reflects revealing insider information in each industry. By combining the policy variables with the probabilities, we obtain the following functional form for η_n :

$$\eta_n = \phi_n \left(\underbrace{(1 - \lambda_n) \frac{x_n^i}{\bar{x}}}_{\text{risk of detection through within-firm information}} + \underbrace{\lambda_n \sum_{k=1}^N \frac{x_{nk}^f}{x_n}}_{\text{risk of detection by invoice cross-checking}} \right) \quad 0 \leq \phi_n \leq 1, \quad 0 \leq \lambda_n \leq 1. \quad (18)$$

When the enforcement expenditure g_n in an industry n goes up, ϕ_n approaches one and the tax administration finds the informal firms in that industry with a higher probability.

4 Market equilibrium and optimal enforcement

Given the two instruments to enforce the VAT, in this part, I find the equilibrium informal production by evaluating the response of informal firms to the enforcement policy. To reduce the risk of unintentional mistakes and whistle-blowing threat, informal firms can control their size. Therefore, they cannot grow as formal firms and their size in this regard depends on the detection possibilities of the tax administration which is reflected in $(1 - \lambda_n)\phi_n x_n^i / \bar{x}$. The second risk of detection is through invoice cross-checking of business transactions. To reduce this risk, informal firms should pay attention to their customers' type. An informal firm has two choices regarding the cross-checking risk: (I) selling the product to all business customers irrespective of their type, (II) not selling to formal customers and match with informal firms.

In action (I), the informal firm sells to both formal and informal clients and may issue a fake invoice for formal customers, resulting in a risk of detection by cross-checking. The advantage of this action is selling the product at the tax inclusive price and earning a money equal to the full VAT payments $e_n = t(p_n x_n^i - \sum_k p_k x_{kn}^i)$. This action can be a desirable strategy when the bulk of customers of the firm are informal and the invoice cross-checking rate λ_n is small in industry n . In comparison, action (II) is matching with an informal client, who cannot apply for input credit, to be insured against the cross-checking risk. However, in this action, in order to persuade the informal client to match, the firm has to share part of the VAT evasion with the buyer. Therefore, if an informal firm in industry n chooses action (II), its gain of evasion is less than e_n .

To characterize the equilibrium after action (II), I assume the informal customers agree to match if they receive a share $\mu \in [0, 1]$ of the seller's gain of matching. Given the form of cost function in (15), when cross-checking is optimal for the administration, the expected loss of the seller per unit of evasion equals to $\theta\phi_n \sum_k x_{nk}^f / x_n$.¹³ Thus, in industry n , while the evasion gain of the informal seller choosing action (I) is e_n , in case it chooses action (II), this amount decreases by factor $\mu e_n \theta\phi_n \sum_k x_{nk}^f / x_n$.

We can find the turning point that informal firms decide between the two actions by comparing the profit in each case. Under action (I), the detection probability η_n is equal

¹³To find this, note that based on the magnitude of the risk that each information source reveals, the detection probability η_n is either increasing or decreasing in λ_n (I show this formally in Proposition 3). In other words, the optimal cross-checking share λ_n for the administration is either zero or one. If $\lambda_n = 0$, the firm's best response is action (I), but if $\lambda_n = 1$ they choose (II) and the equilibrium becomes assortative matching (see Appendix A.5). Thus, when cross-checking is optimal for the administration, the gain of matching per unit of evasion for firm which is $\theta\eta_n$ becomes $\theta\phi_n \sum_k x_{nk}^f / x_n$.

to (18) and overall profit is computed by substituting (18) into (16). Under action (II), the cross-checking risk is removed from the detection probability and we would have $\eta_n = \phi_n(1 - \lambda_n)x_n^i/\bar{x}$. Instead, the matching cost is subtracted from the overall profit (16). With straightforward calculations, we can show that the difference in profits of action (I) and (II) is

$$\pi_n^i(\text{I}) - \pi_n^i(\text{II}) = (\mu - \lambda_n)e_n\theta\phi_n \sum_{k=1}^N \frac{x_{nk}^f}{x_n} \quad (19)$$

If $\lambda_n \leq \mu$, which means the average cost of matching is higher than the risk of cross-checking, the informal firms in industry n prefer action (I). If $\lambda_n > \mu$, informal firms match with informal customers and only supply informal demand. This results in an assortative matching equilibrium similar to De Paula and Scheinkman (2010), in the sense that formal and informal firms trade with the same type shaping two parallel markets.

For illustrative purposes and keep some level of mathematical simplicity, I study assortative matching equilibrium, which lead to the similar results, in Appendix A.5. Here, I just mention that $\lambda_n > \mu$ is optimal for the administration only if market demand is binding for informal firms and they produce below their optimal level. In the following, I assume informal firms produce at their optimum and study the pattern of formality and optimal enforcement across different industries based on their linkages. The below proposition shows the equilibrium relationship between formality defined as $f_n = y_n^f/y_n$, government enforcement, and forward linkages.

Proposition 2 *Formality is positively associated with forward linkages in each industry such that we have*

$$f_n = \tilde{f}_n + \gamma_n \sum_{k=1}^N \alpha_{nk} \frac{y_k}{y_n} f_k \quad \text{and} \quad F = (\mathbf{I} - \mathbf{Y}^{-1}\mathbf{\Gamma}\mathbf{A}\mathbf{Y})^{-1}\tilde{F} \quad (20)$$

where $\tilde{f}_n = 1 - \frac{\bar{x}}{2\theta\phi_n(1-\lambda_n)x_n}$, $\gamma_n = \frac{\min[\lambda_n, \mu]\bar{x}}{2(1-\lambda_n)x_n}$, $F_{(N \times 1)}$ and $\tilde{F}_{(N \times 1)}$ are vectors containing $f_1 \dots f_N$ and $\tilde{f}_1 \dots \tilde{f}_N$ respectively, and $\mathbf{\Gamma}_{(N \times N)}$ is a diagonal matrix consisting $\gamma_1 \dots \gamma_N$ as diagonal elements.

Proof 2 *Appendix A.2*

The effort of the administration to exploit within-firm information to detect informality in industry n is reflected in $\phi_n(1 - \lambda_n)$ and \tilde{f}_n is the formality caused by that. Therefore, the

formula for \tilde{f}_n intuitively indicates that the more effort of tax administration and higher punishment θ reduces the size of informal market.

While \tilde{f}_n is the formalization due to direct enforcement in activity n , the second term reflects the formality transmission by the self-enforcing property. In fact, the particular interest of Proposition 2 is showing how formality in one industry spread to others. From (10), we can see $(\mathbf{I} - \mathbf{Y}^{-1}\mathbf{\Gamma A Y})^{-1}$ is a matrix of forward linkages¹⁴ and by expansion we can write

$$\mathbf{F} = (\mathbf{I} + \mathbf{Y}^{-1}\mathbf{\Gamma A Y} + \mathbf{Y}^{-1}(\mathbf{\Gamma A})^2\mathbf{Y} + \dots)\tilde{\mathbf{F}} \quad (21)$$

where $\mathbf{Y}^{-1}(\mathbf{\Gamma A})^t\mathbf{Y}\tilde{\mathbf{F}}$ is the indirect effect of formality in other industries in rising the formality of an industry through t intermediate goods in between. This equation shows how formality transmit from clients to suppliers under the VAT. If we expand (21) for each element of \mathbf{F}

$$f_n = \underbrace{\tilde{f}_n}_{\text{no linkage effect}} + \overbrace{\sum_{k=1}^N \gamma_n \alpha_{nk} \frac{y_k}{y_n} \tilde{f}_k}^{\text{the firm's own formal costumers}} + \underbrace{\sum_{k=1}^N \sum_{j=1}^N \gamma_n \alpha_{nk} \gamma_k \alpha_{kj} \frac{y_j}{y_n} \tilde{f}_j}_{\text{formal customers through one intermediate good (formal customers of customers)}} + \dots \quad (22)$$

where \tilde{f}_n is formality caused by revealing within-firm information and the next terms gauge the spill-overs of the transmission of formality caused by direct or indirect formal customers of n . Therefore, by utilizing the self-enforcing feature of the VAT, tax administration can increase the formal sector size from \tilde{f}_n to f_n .

Proposition 2 also explains Figure 1. From (8) we have

$$\sum_{k=1}^n \alpha_{nk} \frac{y_k}{y_n} = 1 - \frac{w\beta_n}{y_n} \quad (23)$$

The left hand side of (23) is the direct effect of forward linkages in (22). Thus, we can conclude that the forward linkages of an industry is negatively related to its final demand share.¹⁵ Consequently, an activity with more final demand in total production is more likely

¹⁴The forward linkage index is not completely the same as (10) and has an additional term $\mathbf{\Gamma}$ consisting of γ_n as diagonal elements. The formula of γ_n shows that it is determined by cross-checking share λ_n and is like a policy variable set by the administration. If it is set equal to 1 in all industries, the linkages index becomes equal to (11), otherwise γ_n acts as a fixed weight to the linkages of each industry to compute the linkages. Later, I discuss how the weights are determined at the optimum.

¹⁵ w is the net-of-tax income, and in the presence of a homogeneous VAT, w represents gross household income divided by $1 + t$.

to be informal and has a lower tax registration rate. Note that, using final consumption share only takes into account the direct linkages and not indirect linkages effect from formal customers of customers.

4.1 Optimal Enforcement

After characterizing the formality of each industry, in this section, I derive the optimal enforcement policies to control informality. The objective of the tax administration is maximizing total VAT revenue subtracted by total enforcement expenditure. The VAT revenue for each representative x_n is equal to $t(p_n x_n^f - \sum_{k=1}^N p_k x_{kn}^f)$. By substituting $x_{kn}^f = \alpha_{kn} x_n^f p_n / p_k$ from (3), we obtain

$$t(p_n x_n^f - \sum_{k=1}^N p_k x_{kn}^f) = t(1 - \sum_{k=1}^N \alpha_{kn}) p_n x_n^f = t \alpha_n y_n f_n \quad (24)$$

On the other hand, the administration faces an enforcement expense g_n in each industry which determines the parameter ϕ_n in detection probability. If g_n approaches zero, detection becomes infeasible and if it goes up ϕ_n approaches one. Knowing revenues and costs, we can write the tax administration's optimization problem

$$\max_{\substack{g_1, \dots, g_N \\ \lambda_1, \dots, \lambda_N}} r = \sum_{n=1}^N t \alpha_n y_n f_n - g_n \quad (25)$$

The next two propositions indicates the optimal share of cross-checking and enforcement expenditure in each industry.

Proposition 3 *The optimal share of cross-checking in total enforcement spending of each industry is given by*

$$\lambda_n = \begin{cases} \mu & \text{if } 2 \frac{x_n^i}{\bar{x}} < \sum_k \alpha_{nk} \frac{y_k}{y_n} f_k \\ 0 & \text{if } 2 \frac{x_n^i}{\bar{x}} > \sum_k \alpha_{nk} \frac{y_k}{y_n} f_k \end{cases} \quad (26)$$

Proof 3 *Appendix A.3.*

The two sides of the inequalities are the two sources of information that characterize detection probability in (18). This proposition states that cross-checking is optimal in industry n when it increases detection probability at least twice higher than the firm size threat.¹⁶ In

¹⁶Factor 2 is due to the quadratic functional form of the cost of evasion in x_n^i . If the cost function has higher degree of convexity this number increases and vice-versa.

this case, the tax administration should put effort into cross-checking as much as it does not persuade informal firms to match with their peers. Moreover, the right-hand side of the inequality condition is an indicator of forward linkages – see the second term of (20). Therefore, λ_n is positively associated with forward linkages of an industry.

Proposition 4 *The optimal expenditure for tax enforcement in industry n is positively correlated with its backward linkages and is given by*

$$G^2 = \frac{t\bar{g}\bar{x}}{2\theta} \mathbf{\Lambda}(\mathbf{I} - \mathbf{A}'\mathbf{T})^{-1} \mathbf{A} \quad (27)$$

where G^2 and A are two $N \times 1$ vector with elements g_n^2 and α_n respectively, \mathbf{A} is defined as (9), and $\mathbf{\Lambda}$ is a diagonal matrix comprising $\frac{p_n}{1 - \lambda_n}$.

Proof 4 *Appendix A.4.*

For better illustration, I rewrite (27) in the extensive form for each element

$$g_n = \sqrt{\frac{t\bar{g}\bar{x}p_n}{2\theta(1 - \lambda_n)} \left(\alpha_n + \sum_{k=1}^N \alpha_k \gamma_k \alpha_{kn} + \sum_{k=1}^N \sum_{j=1}^N \alpha_j \gamma_j \alpha_{jk} \gamma_k \alpha_{kn} + \dots \right)} \quad (28)$$

The immediate result of the Proposition 4 is that the optimal expenditure to enforce tax in each industry depends on the square root of the tax rate divided by the punishment multiplier. Indeed, $\sqrt{t\bar{g}\bar{x}/2\theta}$ determines the scale of expenditure the administration has to spend. Moreover, the value of one unit of output p_n and the inverse of the share of within-firm information in enforcement policy $1 - \lambda_n$ increase the optimal enforcement expenditure in the industry.

Apart from these effects, (28) shows that the optimal expenditure is positively associated with backward linkages of the industry defined as $(\mathbf{I} - \mathbf{A}'\mathbf{T})^{-1}$. This reflects the fact that the self-enforcement property of the VAT is from costumers to suppliers not vice-versa. In other words, backwardly linked industries are directly or indirectly customers of forwardly linked ones, so their compliance leads to higher compliance of their suppliers. Therefore, they are better cases for regular tax audit than not backwardly linked industries. This result is of particular interest since it provides a reduce-form expression for optimal enforcement that accounts for formality transmission through input-output linkages.

The backward linkage index used in Proposition 4 is to some extent different from (11), introduced by Rasmussen (1957). The main difference is the appearance of $\gamma_n \propto \lambda_n/(1 - \lambda_n)$

in (28). As shown in Proposition 3, if forward linkages of n is high enough, the optimal policy is $\lambda_n = \mu$, otherwise $\lambda_n = 0$. Therefore, γ_n act as a weight for measuring backward linkages that overstates the linkages with the suppliers that have strong forward linkages, but rules out the linkages with industries that have low forward linkages with $\lambda_n = 0$.

Putting Propositions 4 and 3 together yields the implication of interest: *The tax administration should spend more to reveal within-firm information of backwardly linked industries and make more cross-checking in forwardly linked ones.* For instance, in the simple example of Table 1, the administration should put more effort on the downstream R and try to reveal the within-firm information in this industry (e.g. visiting audit, tempting reward for whistleblowers). But in S, it should dedicate the spending to check the invoices and transactions.

5 Empirical evidence

This section provides evidence about the existence of the VAT self-enforcement and its measurement from the Indian service sector. India is an ideal country for testing this hypothesis given the big size of its shadow economy, large variation in demographics, and policy reforms. The taxes in India are levied by both central and state governments. The central taxes are levied on non-agricultural income and wealth, corporate income and profit, customs duties, excise duties on manufactured products, and services tax. The main state level taxes include states' sales tax, stamp duties and registration fees, motor vehicles tax, agricultural income tax, entertainment tax, profession tax, and electricity duty.

Initiated in 1985 during Rajiv Gandhi's government, Indian taxation system experienced substantial reforms in different directions following to the fiscal crisis and the subsequent liberalization in 1991. The reforms, in a nutshell, led to (i) a substantial reduction in income, profit and wealth tax rates, (ii) a great tariff cut in custom duties, (iii) gradual transformation of central excise duties to the VAT, (iv) service taxes introduction in 1994 and VAT adoption in 2003, (v) changing states' sales tax to state-VAT (in 2005 for the bulk of states), and (vi) administrative improvements. Online Appendix provides the time-line of major tax reforms in India. For more detail on these reforms see Rao and Rao (2006, 2009).

In this paper, I concentrate on service sector of India given service VAT adoption in 2003-04, which provides a natural experiment for testing the link between inter-sectoral linkages and VAT compliance. Service tax in India is levied just by the central (not state) government.

It was introduced in 1994 on insurance and stock brokerages, and overtime gradually more service activities added to the list. The major change happened in 2003-04 by introducing the VAT in service tax and expanding the coverage of the rule to more than 80 activities. After an initial experimentation in 2002-03 for the service inputs, the government budget for 2003-04 made available the credit for taxes paid on inputs to all listed services. Before this reform, the credit for business inputs was not available and service tax – on a number of activities – was charged on purchases by both firms and final consumers. The rate of service tax also has experienced a number of changes. Being 5% initially, the rate increased to 8% in 2003 and then 10.2% in 2004. In the mid-2006, it again rose to 12.24%, and then reduced to 10.3% in 2009.¹⁷

Given the nation-wide service tax policy change in India, I examine the existence of self-enforcing feature by testing whether the VAT has been more effective in reducing the size of shadow economy of forwardly linked activities or not. As shown in Proposition 2, the idea behind this hypothesis is that naturally, firms with strong forward linkages have more business customers and among them formal traders who prefer to buy from VAT compliant suppliers. Thus, after VAT adoption, we expect more tax compliance in forwardly linked activities. For this purpose, I use different data sources; two firm-level surveys on services, and the input-output tables of Indian economy.

5.1 Data

The principal data I use are two consistent surveys on “unorganized service sector” by NSSO¹⁸ which are carried out in 2001-02 (round 57) and 2006-07 (round 63). They have nation-wide representative samples of service enterprises using economic census as the framework and include a variety of information including the status of tax registration. The 2006-07 survey has a specific question about tax registration and whether the firm has tax account number (TAN). In the 2001-02 survey, among other types of registration, the enterprise is asked about states’ sales tax and companies act registration (which includes tax registration). To cover other types of tax, I check whether the tax payments of the enterprise are nonzero or not.¹⁹

¹⁷These facts are obtained from Rao and Rao (2009) and <http://indiataxes.com>.

¹⁸National Sample Survey Organization, which is under Ministry of statistic of India and conducts a variety of surveys including enterprise level on unorganized manufacturing and services. In India, the “organized” sector typically includes enterprises in the public sector and firms registered under the Factories Act, 1948.

¹⁹As an alternative measure, using registration under companies act to define formal status in both surveys leads to virtually identical results. It is good to mention that all NSSO surveys are conducted independent of tax purposes and no information about the identity of the enterprise is supposed to go to the other agencies. As

In both surveys, the sample is highly stratified with a clear distinction between rural and urban areas. The first stage units (FSU) are villages in the rural areas and Urban Frame Survey blocks in the urban areas. The numbers of FSUs are 15,869 and 13,271 in round 57 and 63 respectively. Altogether 244,376 enterprises in 2001-02 and 189,844 enterprises in 2006-07 were surveyed. The services they cover consist of transport, storage, communication, hotels and restaurants, real estate, financial intermediation (just round 63), renting and business activities, education, health, social work, and other community, social and personal service activities. The sample excludes government and public sector undertakings and the units registered under the Factories Act. To have a balanced sample, I drop the financial intermediation services that are just surveyed in 2006.

The second data source, which is used to measure **forward and backward linkages**, is input-output transaction tables of the national accounts statistics of Indian economy published by Central Statistical Organization in 2003-04. The input-output tables consist of 130 activities including 22 service categories²⁰ and include use (commodity \times industry) and make (industry \times commodity) matrices. Using these two, one can construct the (commodity \times commodity) coefficient matrix \mathbf{A} .²¹ Then, similar to (10) and (11), I compute forward and backward linkages in each service sector as

$$FL = \mathbf{Y}^{-1}(\mathbf{I} - \mathbf{A})^{-1}\mathbf{Y}\mathbf{J}, \quad BL = (\mathbf{I} - \mathbf{A}')^{-1}\mathbf{J} \quad (29)$$

Where \mathbf{X} is the diagonal production matrix and \mathbf{J} the summation vector. The final consumption and total demand of each service, used in Figure 1, are drawn from input-output tables too.

it mentioned in the manual, according to the Collection of Statistics Act, violation of any of the confidentiality and secrecy of the information by statistics officer or field staff is prosecuted by or with the consent of appropriate agency.

²⁰Railway transport, Land transport, Water transport, Air transport, auxiliary transport activities, Storage and warehousing, Communication, Trade, Hotels and restaurants, Banking, Insurance, Ownership of dwellings, Education and research, Medical and health, Business services, Computer & related activities, Legal services, Real estate activities, Renting of machinery and equipment, Community, social and personal services, Other services, Public administration.

²¹The methodology for calculating matrix \mathbf{A} from use and make tables are available in the Appendix 2 of the manual of Input-Output Transaction Tables: "Mathematical expression on the methodology of construction of associated matrices", which is available at: http://mospi.nic.in/Mospi_New/upload/report&publication/ftest10/appendix202.pdf

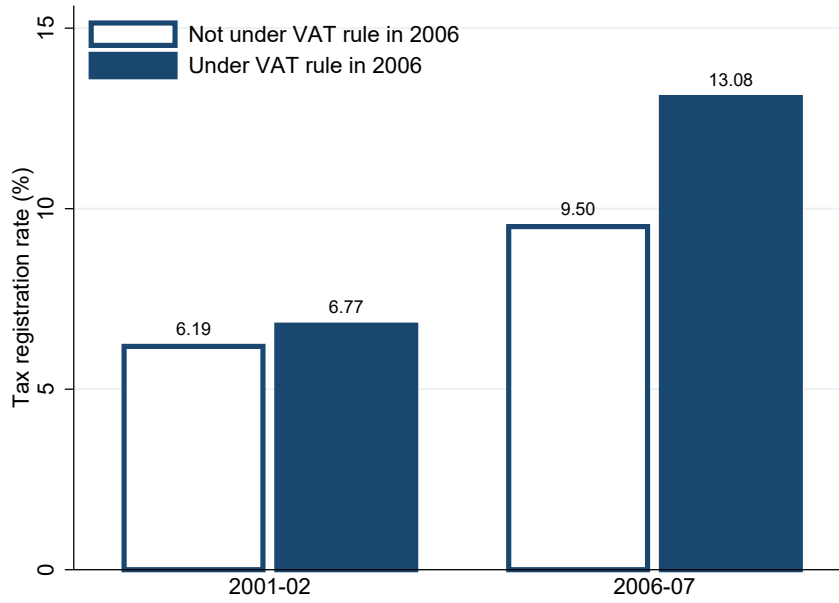


Figure 3: Tax registration rate before and after VAT adoption in 2003. The activities are divided based on the applicability of the service tax rule in 2006. The white bars are the average registration rate in activities that are not listed in service VAT rule until 2006 and the dark bars show the average tax registration rate of activities included in service VAT rule till 2006.

5.2 Overall change in tax compliance of service sector in 2001-2006

Having firms' tax registration status enables me to estimate the size of shadow economy at state-year-activity level using the sample weights. Tax registration status is not specific to the service tax and it refers to all types of registration for central or state government tax agencies. Figure 1 in the introduction, uses the percentage of tax registered firms in each activity – using the sample weights and separately for 2001 and 2006 – in the y-axis. I use the same index in Figure 3 to compares the change in tax registration rate of the services available in the two rounds by decomposing them based on VAT adoption. In particular, the white bars are the average registration rate in activities that are not listed in service VAT rule until 2006 and the dark bars show the average tax registration rate of activities included in service VAT rule till 2006.²² In other words, between 2001 and 2006, the latter was under VAT treatment and the former was as the control group.

In 2001-02, the tax registration rates of the activities excluded and included in service tax

²²The registration rate per year and per activity is available in the online appendix. The classification of services in the NSS surveys are based on 5-digit National Industrial Classification (NIC) code. The number of activities added to the service tax rule till 2006 is 94, but some of them are sub-categories of one 5-digit NIC code. For instance, business auxiliary services, opinion poll, and intellectual property service activities, having a same NIC (74130), are listed as three separate activities in service tax rule. In addition, some other activities under service tax rule are in financial intermediation or trade sectors that are excluded from surveys. The profile of each activity under service tax rule is available at <http://www.servicetax.gov.in/>

rule till 2006 were quite close. On average 6.19% of the enterprises in the first group and 6.77% of the enterprises in the second group were registered under tax administrations. In contrast, these numbers respectively are 9.5% and 13.08% in 2006, meaning around 3% and 6% more registration compared to 2001. To examine the robustness of the measurements, we can compare these numbers with the overall tax collection trend of India. According to World Bank WDI database tax revenue as a percentage of GDP in India grew from 7.94% in 2001 to 11.03% in 2006. This increase coincides with the major reforms in tax administration by the government of India during this period. Notably, launching Tax Information Network, Permanent Account Number (PAN) card service, and Online Tax Accounting System (OLTAS) in January 2004, and extending the legal requirement for filing tax return²³ were some important steps in improving tax administration efficiency in 2001-2006.

In addition to overall growth in tax compliance, Figure 3 shows that between two years before and three years after introduction of the VAT in service tax, the tax registration rate of the activities that the service tax rule applied to grew 3.2% more than the rest. Although administrative advantages and self-enforcing feature of the VAT can be one reason for this evidence, it can also be because of higher effort of the tax administration in the activities under service tax rule. In fact, the adoption of service tax in an activity increases its tax capacity for the administration because it is subject to both direct (income and profit tax) and indirect (service tax) taxes. In contrast, in the other activities that only direct taxes apply the potential tax revenue per firm is relatively less.

5.3 Methodology and summary statistics

To remove the tax effort bias due to service tax adoption from the sample, I drop all services that were not under service tax rule in 2006. In this way, I also eliminate any potential selection bias for activities that are chosen by the government for service tax adoption. Furthermore, because the model characterizes the consequences of VAT adoption – not a change in tax regime from a cascading tax to a VAT – I drop the activities that were added to the service tax rule schedule before 2001. Indeed, without the input refund, service tax is charged at multiple stages of value-added network and “cascades” down to the final stage. The removal of the cascading distortions, may alter the model’s predictions and complicate the interpretation

²³All individuals living in large cities who satisfy at least one of the conditions (ownership of house, cars, membership of a club, ownership of credit card, foreign travel, and subscriber of a telephone connection) became required to file their tax return

of results. Hence in below estimations, the sample only includes activities that were under service tax rule in 2006 but not in 2001. Finally, hair dressing and other beauty treatment services (NIC code 93020) is dropped, because service tax is applied only to beauty treatment not hair dressing and decomposition is not possible. After applying these data cleaning steps, 43 activities remain in the sample which are described in appendix Table A.2.

In order to measure formality, I use two indices: First, **tax registration rate** is the percentage of firms registered under tax authorities (the same index used in Figure 1 and Figure 3) within each state, year, and activity. Second, I define formal and informal production as the level of production that is produced by firms registered and not registered under tax respectively. Then, the **formal production share** refers to the percentage of total production by firms that are registered for central or state tax agencies within each state, year, and activity.

Table 2 shows the summary statistics of formality indices for 2001-02 and 2006-07. Tax registration rate grew from 6.59% in 2001 to 10.81% in 2006 with an average of 8.71 percent. Moreover, while in the whole sample 15.63% of production is produced by formal sector, formality increased from 11.27% in 2001 to 19.95% in 2006.²⁴ Appendix Tables A.1 and A.2 provide the average formality across state and activities respectively. We note a large variation in formality across states and activities. Tax registration rate varies from less than 2% in Bihar and Daman and Diu to over 20% in Arunachal Pradesh and Nagaland. Furthermore, while 40% of service production is created by registered firms in Maharashtra, less than 2% of production is formal in Daman and Diu and Dadra and Nagar Haveli. Looking into activities, we observe less than one percent of firms working as outdoor caterer, fashion designer, and sound recording services are registered for tax, whereas more than 90% are registered in air transport services. Also, almost all of production in cargo handling in water transport is formal, but this number is less than one percent for health club service.

Table 2 also presents the summary statistics of linkages measures that are used in estimations. Forward and backward linkages indicators are on average 1.87 and 1.74 respectively, the correlation between them is 0.26. Appendix Table A.2, presents the variation in the linkages across activities. Storage and warehousing (NIC 2-digit code 63), telecommunication services (NIC 2-digit code 64), and business services (NIC 2-digit code 74) are the most forwardly linked services. In comparison, catering, road transport and storage and warehousing are the

²⁴Note that these numbers are estimated for the sample of “unorganized service sector” as described above and does not cover the whole service sector of India.

most backwardly linked activities. Among all surveyed services, training and coaching (NIC 2-digit code 80) has the lowest linkages with other activities.

The last two variables in Table 2 are the average number of full-time workers per enterprise and total number of firms in thousand in each state and activity. The former captures the effect of firm size and third party reporting by labor, and the latter controls for the fixed firm-level enforcement costs in the estimations. In the whole sample, the number of full-time workers per enterprise and the number of firms per state and activity are on average 2.64 and 4812 respectively. However, service sector enterprises hire more full-time workers in 2006, but the number of firms is bigger in 2001.

Table 2: Summary statistics— The numbers show the summary statistics average across states and activities.

Sample	2001-02		2006-07		The whole sample	
	Mean	SE	Mean	SE	Mean	SE
Tax registration rate	6.587	(15.76)	10.81	(21.96)	8.714	(19.24)
Formal production share	11.76	(23.58)	20.08	(31.98)	15.95	(28.42)
Forward linkages index	1.966	(0.606)	1.770	(0.589)	1.867	(0.605)
Backward linkages index	1.789	(0.407)	1.685	(0.358)	1.737	(0.387)
Average no. workers	2.488	(5.242)	2.786	(5.284)	2.638	(5.263)
Number of firms ('000)	5.106	(21.67)	4.521	(17.74)	4.812	(19.78)

To estimate the effect of service VAT adoption on the relative size of formal sector, I use the following regression setup

$$f_{ist} = a_0 + a_1 VAT_t + a_2 VAT_t \times L_i + a_3 X_{ist} + I_i + S_s \times T_t + \varepsilon_{ist} \quad (30)$$

where f_{ist} is formality index in activity i , state s and time t , VAT_t is equal to one in 2006 and zero in 2001, L_i is one of the linkages indices in activity i , and X_{ist} is a vector of control variables including the average number of workers and total number of firms at activity, state, and year level. In addition, I_i , S_s and T_t are vectors of activity, state, and time fixed effects. By including state \times year fixed effects in (30), I control for all state-level time-variant variables and focus on the differential impact of forward and backward linkages interacted with VAT dummy on formality. In the sample, there are 43 activities and 34 states, but some activities do not exist in all states and the number of observations is 1119 in the regressions.

5.4 VAT adoption, forward linkages and formalization

Table 3 illustrates the estimated coefficients of (30). As a first shot, I do not include the linkages interaction term in column (1) to examine the overall impact of VAT adoption. After controlling for size, firm population, and fixed effect, we find a positive and significant impact of the year dummy of 2006 (post VAT adoption) on formality. Tax registration rate is 3.65% higher in 2006 compared to 2001. Based on the numbers in Figure 3, we can assume half of the increase is due to the other improvements in tax administration. If this is the case, VAT adoption increases tax registration rate by 1.8% which is around one-fifth of 8.71%, the mean of tax registration rate.

The sign of the average number of workers is positive and significant, suggesting higher formal production in bigger firms with more employees. Having one additional worker increases the chance of tax registration by one percent. This is consistent with the third party reporting hypothesis of Kleven et al. (2015) which I used in section 3.1. The effect of total number of firms is negative and significant which can reflect more difficult enforcement in the presence of large number of firms and fixed firm-level administrative costs (Dharmapala et al., 2011). When the number of firms increases by one thousand the chance of tax registration decreases by 0.025%.

Column (2) to (4) of Table 3 presents the results of the estimation of (30) for specifications with only forward linkages, only backward linkages, and both types of linkages. In these estimations, I control for state \times year fixed effects, therefore, the overall improvements between the two years are captured with time dummies. Column (2) shows that activities with higher forward linkages experience more formality after VAT adoption. For example, after VAT adoption in 2003, registration rate of storage and warehousing activities (forward linkages=2.78) increased by 8.97 percentage point more than of training and coaching (forward linkages=1.01).²⁵ However, column (3) implies no significant relationship between tax registration and the interaction term of backward linkages. If we include both linkages indices in one regression, forward linkages remains significant and backward linkages is still insignificant with a sharp reduction in its coefficient (column 4). Similar to column (1), the average number of workers and the number of firms appear with positive and negative signs respectively and their coefficient is significant at 5% level in all specifications.

Column (5) to (8) present the estimation results when the dependent variable is formal

²⁵The economic effect is computed as $(2.78 - 1.01) \times 5.07 = 8.97$

share of production. Column (5) shows that from 2001 to 2006, formal production share increased by 7.74%. If half of the increase is explained by VAT adoption, this number suggest that the impact of VAT adoption on formal production share is 3.86%. In columns (6)-(8), we obtain the same results for the impact of linkages on formal production share. Column (6) suggests that the interaction term of VAT adoption and forward linkages is positively and significantly associated with the dependent variable. The estimated coefficient of backward linkages interacted with VAT adoption is however not significant (column 7). When both of the interaction terms are included in the regression, only forward linkages is positive and significant at 10% level.

To examine the robustness of the results, I conduct a number of additional tests that are available at the online appendix. First, I change definition of formality to registration under companies act – tax registration is its prerequisite – and virtually obtain the identical results. Second, I recompute the linkages measures using the input-output tables of 2006 and observe no change in estimation results.²⁶ Finally, I include the sample of activities under service tax rule before 2001 to increase the number of activities in the sample and find that the impact of interaction term of forward linkages remains positive and significant.

In sum, the empirical evidence confirms a positive and strong impact of forward linkages of an activity on tax registration and formal share of production when the VAT is conducted in the economy. As shown in the model, the self-enforcing feature of the VAT is the main reason for this effect. Hence, the degree of forward linkages of an activity – which can be easily computed using input-output tables of the economy – provides a novel index for tax administrations to measure VAT self-enforcement and move to a more efficient tax enforcement policy.

6 Conclusion

This paper provides the insight that for optimal VAT enforcement, tax administration should take account of the linkages between different activities. Many developing countries with limited tax capacity have moved toward the VAT and the optimality of this reform crucially depends on VAT collection efficiency. The administrative advantage of the VAT is that in addition to common tools to reveal intra-firm information, it provides a unique third-party

²⁶The input-output transaction tables of India are available for 1993, 1998 for 115 sectors and 2003 and 2006 for 130 sectors.

Table 3: Difference-in-difference regressions for the effect of VAT adoption and linkages on formal production. The estimated regression equation is (30). The dependent variables are tax registration rate and share of production by firms registered for tax at each activity, state, and year. VAT adoption is a dummy equal to one after service VAT adoption in 2003. Forward and backward linkages are estimated using input-output tables of Indian economy in 2003. The standard errors are in parentheses and clustered at state level. The * and ** show significance at 10% and 5% respectively.

	Tax registration rate			Formal production share				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VAT adoption dummy (year=2006)	3.652*** (1.190)	-8.521** (3.344)	-3.993 (5.285)	-4.966 (5.196)	7.735*** (2.256)	-7.670 (4.901)	-6.238 (7.145)	-7.046 (7.104)
VAT adoption \times Forward linkages		5.070*** (1.622)		6.695*** (2.179)		5.274** (2.423)		5.559* (3.144)
VAT adoption \times Backward linkages			3.027 (2.931)	-3.828 (3.899)			5.020 (3.943)	-0.672 (5.085)
Average no. workers	0.996*** (0.105)	0.995*** (0.106)	0.988*** (0.108)	0.998*** (0.106)	1.001*** (0.202)	0.990*** (0.201)	0.982*** (0.202)	0.990*** (0.202)
Number of firms ('000)	-0.025*** (0.011)	-0.026*** (0.011)	-0.026*** (0.011)	-0.026*** (0.011)	-0.077*** (0.022)	-0.080*** (0.024)	-0.079*** (0.024)	-0.080*** (0.024)
Observations	1119	1119	1119	1119	1119	1119	1119	1119
R-squared	0.376	0.402	0.397	0.403	0.321	0.356	0.354	0.356
Activity fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State \times year fixed effects	No	Yes	Yes	Yes	No	Yes	Yes	Yes

reported information source via its invoice system. This intrinsic feature causes formality to transmit to forwardly linked activities, because their evasion depends on their customers' informality. Consequently, to formulate the enforcement, tax administration must take into account the spillovers of transmission of formality to the suppliers. For this reason, more enforcement spending should be dedicated to the backwardly linked firms to reveal their internal information. In comparison, tax administrations should spend its resources in forwardly linked activities to cross-check invoices with the corresponding credit claims. Empirical evidence from Indian economy shows that the self-enforcing feature among forwardly linked activities increases formal share of production.

However, the underlying model has some limitations. Generally, this paper neglects some other VAT enforcement issues, as stressed in Keen and Smith (2006), like under-reported sales and over-reported purchases, multiple rates and misclassification of commodities, self-consumption and carousel fraud. In an accompanying paper (Hoseini, 2014), the misreporting in the VAT is studied in detail, but international aspect of VAT fraud like missing trader (carousel) fraud, need more attention of theorists. In spite of these limitations, the findings of this paper have a novel implication for the VAT enforcement design, specially for countries suffering from big size of shadow economy.

References

- Allingham, M. G. and Sandmo, A. (1972). Income tax evasion: a theoretical analysis. *Journal of Public Economics*, 1(3-4):323–338.
- Best, M., Brockmeyer, A., Kleven, H., Spinnewijn, J., and Waseem, M. (2016). Production vs revenue efficiency with limited tax capacity: Theory and evidence from pakistan. *Journal of Political economy*.
- Boadway, R. and Sato, M. (2009). Optimal tax design and enforcement with an informal sector. *American Economic Journal: Economic Policy*, 1(1):27.
- De Paula, u. and Scheinkman, J. A. (2010). Value-added taxes, chain effects, and informality. *American Economic Journal: Macroeconomics*, 2(4):195–221.
- Dharmapala, D., Slemrod, J., and Wilson, J. D. (2011). Tax policy and the missing middle:

- Optimal tax remittance with firm-level administrative costs. *Journal of Public Economics*, 95(9-10):1036 – 1047. Special Issue: The Role of Firms in Tax Systems.
- Ebrill, L. P., Keen, M., Bodin, J.-P., and Summers, V. P. (2001). *The Modern VAT*, volume 1. International Monetary Fund.
- Emran, M. and Stiglitz, J. (2005). On selective indirect tax reform in developing countries. *Journal of Public Economics*, 89(4):599–623.
- Fortin, B., Marceau, N., and Savard, L. (1997). Taxation, wage controls and the informal sector. *Journal of Public Economics*.
- Gordon, R. and Li, W. (2009). Tax structures in developing countries: Many puzzles and a possible explanation. *Journal of Public Economics*, 93(7-8):855–866.
- Hoseini, M. (2014). Misreporting in the value-added tax and the optimal enforcement policy. *Tilburg University, mimeo*.
- Keen, M. (2008). VAT, tariffs, and withholding: Border taxes and informality in developing countries. *Journal of Public Economics*, 92(10-11):1892–1906.
- Keen, M. (2013). The anatomy of the VAT. *IMF Working Paper*, (13-111).
- Keen, M. and Mintz, J. (2004). The optimal threshold for a value-added tax. *Journal of Public Economics*, 88(3):559–576.
- Keen, M. and Smith, S. (2006). VAT fraud and evasion: What do we know, and what can be done? *National Tax Journal*, LIX(4).
- Kleven, H. J., Kreiner, C. T., and Saez, E. (2015). Why can modern governments tax so much? an agency model of firms as fiscal intermediaries. Working Paper 15218, National Bureau of Economic Research.
- Kopczuk, W. and Slemrod, J. (2006). Putting firms into optimal tax theory. *American Economic Review: Papers and Proceedings*, 96(2):130–134.
- Piggott, J. and Whalley, J. (2001). VAT base broadening, self supply, and the informal sector. *American Economic Review*, 91(4):1084–1094.

- Pomeranz, D. (2015). No taxation without information: Deterrence and self-enforcement in the value added tax. *American Economic Review*, 105(8):2539–69.
- Rao, M. G. and Rao, R. K. (2006). Trends and issues in tax policy and reform in india. *India Policy Forum 2005-06*, pages 55–122.
- Rao, M. G. and Rao, R. K. (2009). Tax system reform in india. *Initiative for Policy Dialogue Working Paper Series*.
- Rasmussen, N. (1957). Studies in inter-sectoral relations. *Revue conomique*, 8(6):1103–1104.
- Rauch, J. E. (1991). Modelling the informal sector formally. *Journal of development Economics*, 35(1):3347.
- Schneider, F., Buehn, A., and Montenegro, C. E. (2011). Shadow economies all over the world: New estimates for 162 countries from 1999 to 2007. *Handbook on the shadow economy*, pages 9–77.
- Slemrod, J. and Yitzhaki, S. (2002). Tax avoidance, evasion, and administration. *Handbook of public economics*, 3:1423–1470.
- Stiglitz, J. E. (2009). Development-oriented tax policy. *Initiative for Policy Dialogue Working Paper Series*.
- Tait, A. (1991). *Value-added Tax: Administrative and Policy Issues*. International Monetary Fund.
- Tait, A. A. (1972). *Value added tax*. McGraw-Hill.
- Ten Raa, T. (2006). *The economics of input-output analysis*. Cambridge University Press.

A Appendix

A.1 Proof of Proposition 1

According to the Table 1, the amount of evasion e_n of a representative informal firm in industry n is equal to $e_n = t(p_n x_n^i - \sum_{k=1}^N p_k x_{kn}^i)$. Hence, we can rewrite (16) as

$$\pi_n^i = (1 + t - t\theta\eta_n) \left(p_n x_n^i - \sum_{k=1}^N p_k x_{kn}^i \right) - v_n^i, \quad x_n^i \geq 0. \quad (31)$$

By defining $\tau_n = t(1 - \theta\eta_n)$, we can write (31) as

$$\pi_n^i = (1 + \tau_n) \frac{e_n}{t} - v_n^i \quad (32)$$

which gives the F.O.C as

$$\frac{\partial \pi_n^i}{\partial v_n^i} = \left(\frac{\partial \tau_n}{\partial x_n^i} \frac{e_n}{t} + (1 + \tau_n) p_n \right) \frac{\partial x_n^i}{\partial v_n^i} - 1 = 0 \quad (33)$$

$$\frac{\partial \pi_n^i}{\partial x_{kn}^i} = \left(\frac{\partial \tau_n}{\partial x_n^i} \frac{e_n}{t} n + (1 + \tau_n) p_n \right) \frac{\partial x_n^i}{\partial x_{kn}^i} - (1 + \tau_n) p_k = 0 \quad (34)$$

On the other hand, from (1), we have

$$\frac{\partial x_n^i}{\partial v_n^i} = \alpha_n \frac{x_n^i}{v_n^i}, \quad \frac{\partial x_n^i}{\partial x_{kn}^i} = \alpha_{kn} \frac{x_n^i}{x_{kn}^i} \quad (35)$$

Therefore, we obtain

$$v_n^i = \alpha_n x_n^i \left(\frac{\partial \tau_n}{\partial x_n^i} \frac{e_n}{t} + (1 + \tau_n) p_n \right), \quad x_{kn}^i = \frac{\alpha_{kn} x_n^i}{(1 + \tau_n) p_k} \left(\frac{\partial \tau_n}{\partial x_n^i} \frac{e_n}{t} + (1 + \tau_n) p_n \right) \quad (36)$$

Next, by substituting v_n^i and x_{kn}^i from (36) into (1) and taking (4) into account, it turns out

$$\frac{\partial \tau_n}{\partial x_n^i} \frac{e_n}{t} + (1 + \tau_n) p_n = p_n (1 + \tau_n)^{1-\alpha_n} \quad (37)$$

Thus, at the optimum, firm choices are

$$v_n^i = \alpha_n p_n x_n^i (1 + \tau_n)^{1-\alpha_n}, \quad x_{kn}^i = \frac{\alpha_{kn} p_n x_n^i}{p_k (1 + \tau_n)^{\alpha_n}} \quad (38)$$

By substituting x_{kn}^i from (38) into $e_n = t(p_n x_n^i - \sum_{k=1}^N p_k x_{kn}^i)$, we have

$$\frac{e_n}{t} = p_n x_n^i \left(1 - \frac{\sum_k \alpha_{kn}}{(1 + \tau_n)^{\alpha_n}}\right) = p_n x_n^i \left(1 - \frac{1 - \alpha_n}{(1 + \tau_n)^{\alpha_n}}\right) \quad (39)$$

and by substituting (39) into (37) and rearranging the expression, it turns out

$$\frac{\partial \tau_n}{\partial x_n^i} x_n^i = -(1 + \tau_n) \frac{(1 + \tau_n)^{\alpha_n} - 1}{(1 + \tau_n)^{\alpha_n} - 1 + \alpha_n} \quad (40)$$

Since, $0 \leq \tau_n = t(1 - \theta\eta_n) < 1$, and τ_n is small, I am able to simplify (40) by using the linear approximation of $(1 + \tau_n)^{\alpha_n}$ as $1 + \alpha_n \tau_n$.²⁷ By doing this, we get

$$\frac{\partial \tau_n}{\partial x_n^i} x_n^i = -\tau_n \quad (41)$$

Finally, by substituting $\tau_n = t(1 - \theta\eta_n)$ in (41), we obtain

$$\theta \frac{\partial \eta_n}{\partial x_n^i} x_n^i = 1 - \theta\eta_n \quad (42)$$

A.2 Proof of Proposition 2

From Proposition 1, at the informal firms' optimum, we have $1/\theta = \eta_n + x_n^i \partial \eta_n / \partial x_n^i$. If $\mu > \lambda_n$ and action (I) is chosen and substituting (18) into this yields

$$\frac{1}{\theta \phi_n} = 2(1 - \lambda_n) \frac{x_n^i}{\bar{x}} + \lambda_n \sum_{k=1}^N \frac{x_{nk}^f}{x_n} \quad (43)$$

Because $x_n^f = x_n f_k$ and we have $x_{nk}^f = \alpha_{nk} \frac{p_k}{p_n} x_k^f$ (see (2)–(3)), we can write

$$\frac{x_{nk}^f}{x_n} = \alpha_{nk} \frac{y_k}{y_n} f_k \quad (44)$$

Moreover, at the equilibrium, we have $x_n^i/x_n = y_n^i/y_n = 1 - f_n$. Therefore, by rearranging (43) we get

$$f_n = 1 - \frac{\bar{x}}{2\theta \phi_n (1 - \lambda_n) x_n} + \frac{\lambda_n \bar{x}}{2(1 - \lambda_n) x_n} \sum_{k=1}^N \alpha_{nk} \frac{y_k}{y_n} f_k \quad (45)$$

²⁷Notice that the tax rate is between 0 and 1 (around 0.2 for many countries) and when it multiplies by $0 < 1 - \theta\eta_n < 1$, it becomes even smaller. Therefore, the higher powers of τ_n are negligible for the approximation. For instance, if $\tau_n = 0.1$ and $\alpha_n = 0.5$, the true value and the first order approximation become 1.0488 and 1.05, respectively, with only 0.1% error.

If $\mu < \lambda_n$ and action (II) is chosen, then cross-checking risk is removed from the detection probability and we have $\eta_n = \phi_n(1 - \lambda_n)x_n^i/\bar{x}$, but $\mu e_n \theta \phi_n \sum_k x_{nk}^f/x_n$ is subtracted from the profit of informal firms. By applying both of these changes the profit of a representative informal firm can be written in the same form as (32) with a change in the definition of τ_n such that

$$\tau_n = t(1 - \theta\eta_n - \mu\theta\phi_n \sum_{k=1}^N \frac{x_{nk}^f}{x_n}) \quad (46)$$

Then, we can use (41) to find the optimal decision of the firm.

$$\theta \frac{\partial \eta_n}{\partial x_n^i} x_n^i = 1 - \theta\eta_n - \mu\theta\phi_n \sum_{k=1}^N \frac{x_{nk}^f}{x_n} \quad (47)$$

The equilibrium formality index in each industry can be obtained similar to (43) – (45)

$$f_n = 1 - \frac{\bar{x}}{2\theta\phi_n(1 - \lambda_n)x_n} + \frac{\mu\bar{x}}{2(1 - \lambda_n)x_n} \sum_{k=1}^N \alpha_{nk} \frac{y_k}{y_n} f_k \quad (48)$$

Therefore, by defining $\gamma_n = \frac{\min[\lambda_n, \mu]\bar{x}}{2(1 - \lambda_n)x_n}$, the formality index becomes

$$f_n = 1 - \frac{\bar{x}}{2\theta\phi_n(1 - \lambda_n)x_n} + \gamma_n \sum_{k=1}^N \alpha_{nk} \frac{y_k}{y_n} f_k \quad (49)$$

Next, I employ vectors F and \tilde{F} with dimensionality $N \times 1$ containing f_n and $1 - \frac{\bar{x}}{2\theta\phi_n(1 - \lambda_n)x_n}$, respectively; and $\mathbf{\Gamma}$ as a diagonal $N \times N$ matrices with γ_i as diagonal elements. Then, (49) is written as

$$F = \tilde{F} + \mathbf{Y}^{-1}\mathbf{\Gamma}\mathbf{A}\mathbf{Y}F \quad \Rightarrow \quad F = (\mathbf{I} - \mathbf{Y}^{-1}\mathbf{\Gamma}\mathbf{A}\mathbf{Y})^{-1}\tilde{F} \quad (50)$$

Since $\gamma_n \geq 0$, according to Ten Raa (2006, p. 22, Theorem 2.2) (multiply the column vector by $\mathbf{\Gamma}$), F always has a unique solution.

A.3 Proof of Proposition 3

From Proposition 2, when $\mu > \lambda_n$, at equilibrium we have

$$x_n^f = x_n - \frac{\bar{x}}{2\theta\phi_n(1-\lambda_n)} + \frac{\lambda_n\bar{x}}{2(1-\lambda_n)} \sum_{k=1}^N \frac{x_{nk}^f}{x_n} \quad (51)$$

Since $x_n^i = x_n - x_n^f$, by rearranging (2), we can write

$$2(1-\lambda_n)\frac{x_n^i}{\bar{x}} + \lambda_n \sum_{k=1}^N \frac{x_{nk}^f}{x_n} = \frac{1}{\theta\phi_n} \quad (52)$$

Then, the derivative of x_n^i with respect to λ_n can be derived from (52)

$$2(1-\lambda_n)\frac{1}{\bar{x}}\frac{\partial x_n^i}{\partial \lambda_n} - 2\frac{x_n^i}{\bar{x}} + \sum_{k=1}^N \frac{x_{nk}^f}{x_n} = 0, \quad (53)$$

$$\frac{\partial x_n^i}{\partial \lambda_n} = \frac{\bar{x}}{2(1-\lambda_n)} \left(2\frac{x_n^i}{\bar{x}} - \sum_{k=1}^N \frac{x_{nk}^f}{x_n} \right) \quad (54)$$

Because $1-\lambda_n > 0$, x_n^i is monotone in λ_n and the sign of $\frac{\partial x_n^i}{\partial \lambda_n}$ depends on $2x_n^i/\bar{x} - \sum_k x_{nk}^f/x_n$. If $2x_n^i/\bar{x} > \sum_k x_{nk}^f/x_n$, then x_n^i is increasing in λ_n and the best choice to reduce informality is $\lambda_n = 0$. On the other hand, if $2x_n^i/\bar{x} < \sum_k x_{nk}^f/x_n$, the derivative is negative and the tax administration should focus on cross checking as much as possible ($\lambda_n = 1$). However, according to Proposition 2, if λ_n increases such that $\lambda_n > \mu$, the informal firm start matching with their peers to eliminate the cross-checking risk from their detection probability. The turning point that makes an informal firm indifferent between matching and not matching is $\lambda_n = \mu$. Hence, if $\lambda_n > \mu$ the administration gains nothing from increasing λ_n – even if $2x_n^i/\bar{x} < \sum_k x_{nk}^f/x_n$ – and its best choice is $\lambda_n = \mu$. In sum, the optimal cross-checking in each industry can be written as

$$\lambda_n = \begin{cases} \mu & \text{if } 2x_n^i/\bar{x} < \sum_k x_{nk}^f/x_n \\ 0 & \text{if } 2x_n^i/\bar{x} > \sum_k x_{nk}^f/x_n \end{cases} \quad (55)$$

Finally, using (44), we obtain (26).

A.4 Proof of Proposition 4

Using (22), I can rewrite (25) as

$$r = \sum_{n=1}^N t\alpha_n y_n \left(\tilde{f}_n + \sum_{k=1}^N \gamma_n \alpha_{nk} \frac{y_k}{y_n} \tilde{f}_k + \sum_{k=1}^N \sum_{j=1}^N \gamma_n \alpha_{nk} \gamma_k \alpha_{kj} \frac{y_j}{y_n} \tilde{f}_j + \dots \right) - g_n \quad (56)$$

which yields

$$r = \sum_{n=1}^N t\alpha_n y_n \tilde{f}_n + \sum_{n=1}^N \sum_{k=1}^N t\alpha_n \gamma_n \alpha_{nk} y_k \tilde{f}_k + \sum_{n=1}^N \sum_{k=1}^N \sum_{j=1}^N t\alpha_n \gamma_n \alpha_{nk} \gamma_k \alpha_{kj} y_j \tilde{f}_j + \dots - \sum_{n=1}^N g_n \quad (57)$$

By re-indexing we obtain

$$r = \sum_{m=1}^N t y_m \tilde{f}_m (\alpha_m + \sum_{k=1}^N \alpha_k \gamma_k \alpha_{km} + \sum_{k=1}^N \sum_{j=1}^N \alpha_k \gamma_k \alpha_{kj} \gamma_j \alpha_{jm} + \dots) - g_m \quad (58)$$

and since $\tilde{f}_m = 1 - \frac{\bar{x}}{2\theta\phi_m(1-\lambda_m)x_m}$ and $\frac{1}{\phi_m} = \frac{\bar{g}}{g_m}$

$$\frac{\partial \tilde{f}_m}{\partial g_m} = \frac{\bar{x}\bar{g}}{2\theta(1-\lambda_m)x_m g_m^2}, \quad (59)$$

γ_n is independent of g_n and the FOC of (58) becomes

$$\frac{\partial r}{\partial g_m} = \frac{t y_m \bar{x} \bar{g}}{2\theta(1-\lambda_m)x_m g_m^2} (\alpha_m + \sum_{k=1}^N \alpha_k \gamma_k \alpha_{km} + \sum_{k=1}^N \sum_{j=1}^N \alpha_k \gamma_k \alpha_{kj} \gamma_j \alpha_{jm} + \dots) - 1 = 0 \quad (60)$$

which result in the optimal expenditure in each industry as

$$g_m^2 = \frac{t\bar{g}\bar{x}p_m}{2\theta(1-\lambda_m)} (\alpha_m + \sum_k \alpha_k \gamma_k \alpha_{km} + \sum_k \sum_j \alpha_k \gamma_k \alpha_{kj} \gamma_j \alpha_{jm} + \dots) \quad (61)$$

and in the matrix form

$$G^2 = \frac{t\bar{g}\bar{x}}{2\theta} \mathbf{\Lambda} (\mathbf{I} - \mathbf{A}'\mathbf{\Gamma})^{-1} \mathbf{A} \quad (62)$$

where G^2 and A are vectors with elements g_n^2 and α_n , respectively; $\mathbf{\Lambda}$ is a diagonal matrix with elements $\frac{p_n}{1-\lambda_n}$ and $\mathbf{\Lambda}$ is defined as (9).

A.5 Market demand constraint and assortative matching equilibrium

In the main model, I assumed that the informal firms are not constrained by market demand and their optimal decisions are always found by the internal optimum of their profit maximization problem. In this part, I want to analyze their behavior when market demand is binding and imposes a corner solution to their decision. As mentioned above, an informal firm can choose between two actions to deal with cross-checking: (I) selling to all types of clients, (II) commit and match with an informal client. In each case, I denote the internal solution of the optimal production of an informal firm (as the case of Proposition 1) by x_n^{i*} and the maximum purchase of available customers for a representative informal firm by \hat{x}_n^i . In a similar way, I define $y_n^{i*} = p_n x_n^{i*}$ and $\hat{y}_n^i = p_n \hat{x}_n^i$. If $y_n^{i*} > \hat{y}_n^i$, the market demand is binding and producing at the level of y_n^{i*} is not possible for informal firms. In order to take market demand into consideration, I assume that final consumers are indifferent between buying from formal and informal markets such that the share of informal sales in final consumption of industry n is $\frac{y_n^i}{y_n} w\beta_n$, where $w\beta_n$ is the value of final consumption in industry n (see (8)).²⁸ Then, we can find the production level and payoff of an informal firm from choosing each action. Given the concavity of the profit, the actual production value of the informal firm after each action will be $y_n^i = \min[y_n^{i*}, \hat{y}_n^i]$. Table 4 shows y_n^{i*} and \hat{y}_n^i after choosing action (I) and (II).

Table 4: Outcomes of an informal firm from action (I) (selling to all types of firms) and action (II) (not selling to formal firms) – y_n^{i*} represents the internal solution of optimization problem, while \hat{y}_n^i is the market available for a representative informal firm.

	y_n^{i*}	\hat{y}_n^i
(I)	$\frac{\bar{x}}{2(1-\lambda_n)x_n} \left(\frac{y_n}{\theta\phi_n} - \lambda_n \sum_k \alpha_{nk} y_k^f \right)$	$\frac{y_n^i}{y_n} w\beta_n + \sum_k \alpha_{nk} y_k^i + \sum_k \alpha_{nk} y_k^f$
(II)	$\frac{\bar{x}}{2(1-\lambda_n)x_n} \left(\frac{y_n}{\theta\phi_n} - \mu \sum_k \alpha_{nk} y_k^f \right)$	$\frac{y_n^i}{y_n} w\beta_n + \sum_k \alpha_{nk} y_k^i$

Moving from (I) to (II), the value of market demand \hat{y}_n^i drops, but when cross-checking is optimal and $\lambda_n > 0$ the endogenous optimum y_n^{i*} increases. The market available after action (I) is the total demand in the industry and it is the optimal choice of informal firms when $\lambda_n = 0$. On the other hand, when cross-checking is optimal and action (I) has lower y_n^{i*} , informal suppliers, missing the formal market demand, match with their informal clients.

²⁸alternatively, one can assume a fixed share of final consumers for each type of firms, but as long as the inter-firm factors does not affect the final demand of each group the qualitative results are the same.

According to Table 4, the equilibrium point may change based on whether \hat{y}_n^i is binding or not. When there is no restriction from market demand, the results are presented in Proposition 4 and 3. In comparison, when the market demand is binding, FOC of profit maximization in proposition 1 and 2 are not valid anymore and informal production in each industry is determined by the production of the clients. There are two equilibriums in which the market demand \hat{y}_n is binding. First, if $y_n^{i*}(\text{I}) > \hat{y}_n^i(\text{I})$, informal firms choose (I) and supply all market demand in industry n . In reality, this can happen when market demand for product n is so small that the tax administration leaves it out from taxation (e.g. exemption policies). The second case is when $y_n^{i*}(\text{II}) > \hat{y}_n^i(\text{II})$ and the optimal λ_n is positive. This equilibrium is an assortative matching equilibrium, in the sense that informal firms match with the same type customers and supply all of the informal demand while, formal firms just sell in the formal market to formal businesses or final consumers. This equilibrium is in close relation to the results of De Paula and Scheinkman (2010) which show under VAT each firm trades with the same type. The below Proposition presents the optimal policy when the equilibrium in some industries are AM.

Proposition 5 *If AM holds in industries indexed by $M + 1, \dots, N$, then the optimal policies are*

1. If $M < n \leq N$ (n is under AM): $\lambda_n = 1$ and $g_n = \frac{\bar{g}x_n}{\theta \sum_{k=1}^N x_{nk}^f}$.
2. If $1 \leq n \leq M$ (n is not under AM): λ_n is found from Proposition 3 and

$$G_1^2 = \frac{t\bar{g}\bar{x}}{2\theta} \mathbf{\Lambda}_1 \left(\mathbf{I} - \mathbf{\Gamma}_1 \mathbf{A}'_{11} - \mathbf{D}' \mathbf{\Gamma}_2 \mathbf{A}'_{12} \right)^{-1} (\mathbf{A}_1 + \mathbf{D} \mathbf{A}_2) \quad (63)$$

where G_1^2 is the vector containing g_1^2, \dots, g_M^2 of the M industries that are not under AM, subscripts 1 and 2 represent the decomposition of the corresponding vector or matrix into $n \leq M$ and $n > M$ respectively, such as $\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \end{bmatrix}$, $\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}$, etc., and $\mathbf{D} = (\mathbf{I} - \mathbf{Y}_2^{-1} \mathbf{B} - \mathbf{A}_{22})^{-1} \mathbf{A}_{21}$.

Proof 5 *Online appendix.*

Here, we have two groups of firms: $1 \leq m \leq M$ in which AM is not the equilibrium and $M < k \leq N$ which are under AM. The results 1 indicates that in the industries that are

under AM, the administration should focus only on invoice cross-checking. In this case, the level of expenditures are such that $\pi_k^{i*}(\text{I}) \leq \hat{\pi}_k^i(\text{II})$.

On the other hand, according to the result 2, the optimal expenditures for industries $1 \leq m \leq M$ are similar to Proposition 4, but with some changes due to AM. Basically, AM imposes a linear relationship between informality in industries $M < k \leq N$ and $1 \leq m \leq M$ which is reflected in a transformation matrix \mathbf{D} in the sense that $X_2^i = \mathbf{D}X_1^i$. This relationship transforms X_2^i into X_1^i and reduces the dimension of optimization matrix from N to M . Therefore, in comparison to Proposition 4, here $\mathbf{\Gamma}\mathbf{A}'$ and A (both with N rows) are replaced by $\mathbf{\Gamma}_1\mathbf{A}'_{11} + \mathbf{D}'\mathbf{\Gamma}_2\mathbf{A}'_{12}$ and $A_1 + \mathbf{D}A_2$ (both with M rows) respectively. Hence, if binding constraints from demand side is added to the model, still the results of Proposition 4 and 3 are valid with the difference that here X_2^i is indirectly affected by G_1 through X_1^i but not by G_2 . The role of G_2 is just to prevent firms from choosing action (I).

When AM is not the initial equilibrium, tax administration may be able to shift the equilibrium to the point that AM holds. According to Proposition 5, AM can be imposed to be the equilibrium in industry n by setting $\lambda_n = 1$ and $g_n^{AM} = \hat{g}x_n / \sum_k \theta x_{nk}^f$. Two factors can make AM the desirable equilibrium for the administration. First, under AM, informal firms constrained by the market demand, produce below their optimal level which leads to less informality. Second, if formal firms comprise a large share of clients of industry n (large $\sum_k x_{nk}^f$), the administration can spend much less in industry n to hold AM (small g_n^{AM}). To increase the size of formal customers $\sum_k x_{nk}^f$, the administration can increase enforcement in the industries that buy product n as input. As an example, if industry m is the only customer of industry n , the administration can impose AM in n by increasing g_m and choosing $\lambda_n = 1$. Then informality in both m and n decline with the possibility that g_n under AM is less than before. Thus, imposing AM in industry n can be optimal if the benefits (higher f_n, f_m and possibly less g_n) beats the cost (higher g_m). For finding the industries that AM is the desirable equilibrium, the administration needs to compare the overall benefits and costs with and without imposing AM in each industry which can be very complicated to model analytically. However, looking at the denominator of g_n^{AM} , we can say that the required expenditure to impose AM in industries that have low forward linkages is large and if all client of an industry are final consumers, g_n^{AM} approaches infinity and AM cannot be implemented. Therefore, AM is more likely to be the optimal equilibrium in an industry with strong forward linkages and large number of formal customers.

Table A.1: Formality estimates for Indian states – Tax registration rate (TRR) is the percentage of firm registered under tax. Formal share of production (FSP) measures the percentage of production by tax registered firms. Both estimates are averaged across services activities as listed in Table A.2. The indices are estimated by the author using representative NSSO surveys.

state	TRR	FSP	state	TRR	FSP
Jammu and Kashmir	2.63	3.71	Assam	6.95	9.43
Himachal Pradesh	7.45	12.81	West Bengal	13.69	29.79
Punjab	2.97	9.46	Jharkhand	6.79	14.28
Chandigarh	8.97	14.75	Orissa	6.65	12.80
Uttarkhand	2.88	3.08	Chattisgarh	5.26	5.93
Haryana	5.12	10.88	Madhya Pradesh	10.00	16.73
Delhi	12.50	20.60	Gujarat	10.22	28.55
Rajasthan	7.59	10.82	Daman and Diu	1.25	1.28
Uttar Pradesh	8.75	15.13	Dadra and Nagar Haveli	2.08	1.93
Bihar	1.99	6.61	Maharashtra	15.89	39.99
Sikkim	3.07	3.06	Andhra Pradesh	12.78	21.22
Arunachal Pradesh	20.47	20.98	Karnataka	10.63	18.59
Nagaland	20.98	22.80	Goa	6.36	6.37
Manipur	13.11	19.59	Kerala	7.35	18.56
Mizoram	13.15	11.27	Tamil Nadu	5.39	10.79
Tripura	6.71	7.91	Puducherry	5.51	10.32
Meghalaya	19.86	23.99	Andaman and Nicobar Islands	3.68	5.06

Table A.2: Summary statistics at industry level – Tax registration rate (TRR) is the percentage of firm registered under tax. Formal share of production (FSP) measures the percentage of production by tax registered firms. The sample includes the activities included in services tax rule between 2001 and 2006, and is chosen based on various notifications of service tax rule, Government of India (for details see each service profile at <http://www.servicetax.gov.in>). Services under financial intermediation and trade categories are not sampled in the round 57 of NSS survey and are not listed here. The correlation between forward and backward linkages is 0.26.

NIC	Activity	TRR	FPS	FL	BL
55204	Outdoor Caterer	0.65	4.17	1.39	2.23
60231	Transport of goods by Road	11.73	22.36	1.84	2.22
61100	Port Service: Water transport, inland	45.41	47.45	1.82	1.82
61200	Port Service : Water transport, sea and coastal	1.01	1.29	1.82	1.82
63011	Cargo Handling: Land Transport	24.79	27.30	1.98	1.95
63012	Cargo Handling: Water Transport	84.89	99.98	1.98	1.95
63013	Cargo Handling: Air Transport	91.67	96.39	1.98	1.95
63021	Storage & Warehousing: Agriculture	13.48	23.89	2.78	2.55
63022	Storage & Warehousing: Cold Storage	39.55	54.26	2.78	2.55
63023	Storage & Warehousing: Non-agriculture	11.23	23.55	2.78	2.55
64110	Mailing List Compilation and Mailing	22.22	14.47	2.49	1.53
64202	Internet Telecommunication Service; Broadcasting	12.16	15.76	2.49	1.53
64203	Telecommunication Service: Maintenance	1.20	1.72	2.49	1.53
64204	Cable Operator	14.82	26.93	2.49	1.53
72294	On-line Information: Web-page designing	10.17	25.44	1.20	1.52
72295	Maintenance or Repair: Software	38.10	48.26	1.20	1.52
72400	On-line Information: Database Access or Retrieval Service	11.78	26.35	1.20	1.52
72501	Maintenance or Repair: Computer Hardware	14.25	31.10	1.20	1.52
72502	Maintenance or Repair: Accounting Machinery	12.50	11.26	1.20	1.52
72901	Internet Cafe	15.05	17.85	1.20	1.52
72909	Maintenance or Repair: Other Computer Based Systems	8.35	16.33	1.20	1.52
74130	Business Auxiliary; Opinion Poll; Intellectual Property Service	16.67	18.03	2.57	2.07
74140	Scientific or Technical Consultancy	13.86	29.05	2.57	2.07
74930	Cleaning Service	26.41	36.62	2.57	2.07
74940	Photography	3.96	10.84	2.57	2.07
74950	Packaging Service	5.05	6.98	2.57	2.07
74992	Business supporting activities	5.71	15.20	2.57	2.07
74994	Fashion Designer	0.96	2.09	2.57	2.07
74999	Business Exhibition; Pandal or Shamiana Service	4.81	14.52	2.57	2.07
80902	Training or Coaching: Coaching Centres	5.05	13.91	1.01	1.20
80903	Training or Coaching: by individuals	2.20	6.09	1.01	1.20
80904	Training or Coaching: Computer Courses	3.86	13.13	1.01	1.20
91990	Clubs and Associations	5.84	23.52	1.83	1.66
92111	Video Tape Production: Motion Pictures	39.05	58.15	1.59	1.48
92113	Video Tape Production: Video	5.27	5.65	1.59	1.48
92115	Sound Recording	0.40	8.27	1.59	1.48
92132	T.V. & radio Programme Production Services	11.30	20.99	1.59	1.48
92411	Fitness Centre	3.63	12.01	1.59	1.48
92412	Health Club	1.14	0.75	1.59	1.48
92413	Event Management: Sport	4.60	8.00	1.59	1.48
92490	Convention Centre	4.81	6.98	1.59	1.48
93010	Dry Cleaning	2.62	3.69	1.83	1.66
93030	Event Management: Funeral	2.19	4.56	1.83	1.66

Value-added tax and Shadow economy: The Role of Intersectoral Linkages

Mohammad Hoseini, December 2015

Online Appendix

Proof of Proposition 5

If $\lambda_n = 0$, informal firms are always better off by choosing (I) and AM is not the equilibrium. Thus if AM is the equilibrium $\lambda_n > 0$. Moreover, as shown in Proposition 3, because the profits are monotone in λ_n , its optimal level is either zero or one. Under AM, the administration does not need to make the firm indifferent between $y_n^{i*}(\text{I})$ and $y_n^{i*}(\text{II})$ (since $y_n^{i*}(\text{II}) \geq \bar{y}_n^i(\text{II})$), thus λ_n is not constrained by the corresponding search cost μ and at optimum $\lambda_n = 1$.

On the other hand, holding AM requires $\pi_n^{i*}(\text{I}) \leq \bar{\pi}_n^i(\text{II}) \leq \pi_n^{i*}(\text{II})$. This indicates that for imposing AM in industry n , the administration has to set the efforts such that the informal firms are indifferent between actions (I) and (II) while the market demand is binding. When $\lambda_n = 1$, we always have $\pi_n^{i*}(\text{I}) \leq \pi_n^{i*}(\text{II})$, and under AM the demand constraint yields $\bar{\pi}_n^i(\text{II}) \leq \pi_n^{i*}(\text{II})$, thus AM will be the equilibrium if g_n is chosen such that $\pi_n^{i*}(\text{I}) \leq \bar{\pi}_n^i(\text{II})$. By substituting (38) and (39) in (32), we obtain the profit after action (I) as

$$\pi_n^{i*}(\text{I}) = (1 + \tau_n) \left(1 - \frac{1 - 2\alpha_n}{(1 + \tau_n)^{\alpha_n}} \right) y_n^i \quad (64)$$

where $\tau_n = t(1 - \theta\eta_n)$. With a first order approximation similar to Proposition 1, (64) becomes

$$\pi_n^{i*}(\text{I}) = (1 + \tau_n)^{1 - \alpha_n} (2 + \tau_n) y_n^i \quad (65)$$

When $\lambda_n = 1$, we have $\eta_n = \phi_n \sum_k x_{nk}^f / x_n$ - see (18). This means that τ_n is independent of y_n^i and if $1 + \tau_n > 0$, the optimal production tends to infinity. Therefore, the level of enforcement expenditure g_n that holds $\pi_n^{i*}(\text{I}) \leq \bar{\pi}_n^i(\text{II})$ leads to $1 + \tau_n \leq 0$. The level is equal to

$$g_n = \frac{\bar{g}x_n}{\theta \sum_{k=1}^N x_{nk}^f} \quad (66)$$

In the next step, I want to derive the optimal policy when AM is the equilibrium in some industries. Without loss of generality, I divide the industries into two groups: Industries $1, \dots, M$ that the AM is not the equilibrium, and $M+1, \dots, N$ in which AM is held. If AM equilibrium holds in industry n , we have

$$y_n^i = \sum_{k=1}^N \alpha_{nk} y_k^i + y_n^i \frac{w\beta_n}{y_n} \quad (67)$$

(67) gives no answer for the industries that possess no forward linkages ($\sum_k \alpha_{nk} y_k = 0$ and $y_n = w\beta_n$) since they do not have any business customers. Therefore, the industries with no forward linkage are indexed among $1, \dots, M$. The informal firms that work in industries $1, \dots, M$ get the same payoff as before, and the optimal cross-checking for them can be obtained from Proposition 3. Now, if I decompose vector Y into Y_1 and Y_2 consisting of

y_1, \dots, y_M and y_{M+1}, \dots, y_N respectively, and \mathbf{A} into four corresponding sub-matrices such that $\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}$, we can write (67) as

$$Y_2^i = \mathbf{A}_{21}Y_1^i + \mathbf{A}_{22}Y_2^i + w\mathbf{Y}_2^{-1}\mathbf{B}_2Y_2^i \quad (68)$$

where \mathbf{Y}_2 is a diagonal metrics consisting of y_{M+1}, \dots, y_N respectively. From (68) we can find

$$Y_2^i = (\mathbf{I} - \mathbf{Y}_2^{-1}\mathbf{B}_2 - \mathbf{A}_{22})^{-1}\mathbf{A}_{21}Y_1^i, \quad (69)$$

so by knowing Y_1^i , Y_2^i is obtained from (69). Now, we can use this results to find the optimal policy. Similar to (25) tax administration's objective becomes

$$r = t \sum_{n=1}^N \alpha_n y_n^f - g_n \quad (70)$$

which can be rewritten in matrix form as

$$r = t\mathbf{A}'\mathbf{Y}^f - \mathbf{I}'\mathbf{G} \quad (71)$$

where \mathbf{G} is a vector of g_n s. From (69), define $\mathbf{D} = (\mathbf{I} - \mathbf{Y}_2^{-1}\mathbf{B}_2 - \mathbf{A}_{22})^{-1}\mathbf{A}_{21}$, then

$$Y_2^f = Y_2 - Y_2^i = Y_2 - \mathbf{D}Y_1 + \mathbf{D}Y_1^f \quad (72)$$

Then, we can decompose r for the two groups of industries as

$$r = t(\mathbf{A}'_1 + \mathbf{A}'_2\mathbf{D})Y_1^f + \mathbf{A}'_2(Y_2 - \mathbf{D}Y_1) - \mathbf{I}'\mathbf{G} \quad (73)$$

Define $\mathbf{\Lambda}$ and \mathbf{Y} as diagonal matrices containing $\frac{p_n}{1-\lambda_n}$ and y_n as elements respectively, then we have $\mathbf{Y}^f = \mathbf{Y}F$ and similar to (50) we can find

$$Y_1^f = (Y_1 - \frac{\bar{g}\bar{x}}{2\theta}\mathbf{\Lambda}_1G_1^{-1}) + \mathbf{A}_{11}\mathbf{\Gamma}_1Y_1^f + \mathbf{A}_{12}\mathbf{\Gamma}_2Y_2^f \quad (74)$$

where G_1^{-1} is a $M \times 1$ vector comprising $1/g_1, \dots, 1/g_M$. Using (72), it turns out

$$Y_1^f = \left(\mathbf{I} - \mathbf{A}_{11}\mathbf{\Gamma}_1 - \mathbf{A}_{12}\mathbf{\Gamma}_2\mathbf{D}\right)^{-1} \left(Y_1 - \frac{\bar{g}\bar{x}}{2\theta}\mathbf{\Lambda}_1G_1^{-1} + \mathbf{A}_{12}\mathbf{\Gamma}_2(Y_2 - \mathbf{D}Y_1)\right) \quad (75)$$

Substituting (75) in (73) yields

$$r = r_0 - \frac{t\bar{g}\bar{x}}{2\theta}H\mathbf{G}^{-1} - \mathbf{I}'\mathbf{G} \quad (76)$$

where r_0 is a constant independent of \mathbf{G} , and H is a row vector equal to

$$H = (\mathbf{A}'_1 + \mathbf{A}'_2\mathbf{D})\left(\mathbf{I} - \mathbf{A}_{11}\mathbf{\Gamma}_1 - \mathbf{A}_{12}\mathbf{\Gamma}_2\mathbf{D}\right)^{-1}\mathbf{\Lambda}_1 \quad (77)$$

Thus, if h_m is the m^{th} element of H

$$r = r_0 - \sum_{m \leq M} \frac{t\bar{g}\bar{x}}{2\theta g_m} h_m - g_m - \sum_{k > M} g_k \quad (78)$$

and the FOC for g_m results

$$g_m^2 = \frac{t\bar{g}\bar{x}}{2\theta} h_m \quad (79)$$

which can be rewritten in vector form as

$$G_1^2 = \frac{t\bar{g}\bar{x}}{2\theta} H' \quad (80)$$

Which yields to (63).

Time-line of taxation reforms in India

Tax reforms in India initialized in 1985 during Rajiv Gandhis government and enormous progress has been made since 1991 liberalization. In below, the timeline of main reforms is listed.

Reforms in direct taxes

1. Personal income tax:

- 1973: 11 brackets, rate range 10 to 85%
- 1985: 4 brackets, less than 50%
- 1992: 3 brackets, 20,30, 40%
- 1997: 3 brackets, 10, 20, 30%

2. Corporate income tax:

- Before 1990s: taxing the profits of companies and dividends of shareholders between 45-60%
- In 1993, rate was unified at 40%
- In 1996, MAT (minimum alternative tax) pay 30% on book profit
- In 1997, 35% and 10% on dividend shifted from individuals to companies
- In 2000, dividend tax increased to 20%, then reduced to 10% in 2001 and levied on shareholders
- In 2003, again levied from companies
- Major corporate tax preferences investment and depreciation allowance
- Securities transactions in 2004 and tax of 0.1% on all cash withdrawal above 25000 from current account in 2005
- In 2005, CIT of domestic firms reduced to 30%

3. Wealth tax: 1973: 5%, 1985: 2.5%, 1992: 1% and financial assets excluded

4. From 2004, 2% surcharge in all tax for education

Reforms in indirect tax

1. Central excise duties, moving toward VAT:

- after independence mostly raw materials and intermediate goods,
- 1975: extended to all manufacturing products, but with a complex and distortionary structure,
- 1986: MODVAT was implemented (not covering capital goods),
- 1993: converted to ad valorem rate except few commodities like tea, cement, cigarette,
- 1994: extended to capital goods and petroleum products,
- 1996: MODVAT covered a majority of commodities,

- 1999: 11 rates merged to 3 rates,
 - 2000: 3 rates merged to 16% and called Central VAT,
 - 2003: again 3 rates structure (8, 16, 24%).
2. Custom duties:
- Prior to 1990s: tariff rates were high and complex;
 - weighted average rate: 38% in 1980, 87 in 1989;
 - Peak rates: 400% in 1990, 1991: 150%, 1992:110%, 1993:85%, 2007:10%;
 - Number of major rates was reduced from 22 in 1990 to 4 in 2003.
3. Service taxes:
- entertainment, passengers and goods are assigned to states;
 - In 1994, government levied tax on insurance, stock brokerages, and telecommunication;
 - In 2003 rate increased from 5 to 8% and the list expanded;
 - In 2004 rate was 10% and the list included 80 services.
4. States sales tax:
- Before introduction of VAT, there was no harmony in the rates among the states;
 - A uniform categorical floor rate of sales tax is implemented with effect from 2000;
 - Introduction of VAT in 21 states in 2005 (Haryana in 2003). 8 states assigned later;
 - Urban local bodies impose a tax on the entry of goods (octroi);
 - In designing VAT, attempts made to harmonize it among the state, keeping also a little flexibility;
 - The general rates:
 - 0% or exempted: Agricultural product and basic necessities,
 - 1%: precious metals,
 - 4% (5% after 2009): common consumption goods,
 - 12.5%: the rest,
 - 20% and above: Special items like Fuel and petrol.
5. Next step in indirect tax reform is ongoing: Goods and Services Tax (GST) - proposed to be introduced by April 2016.

Tax registration rate (%)

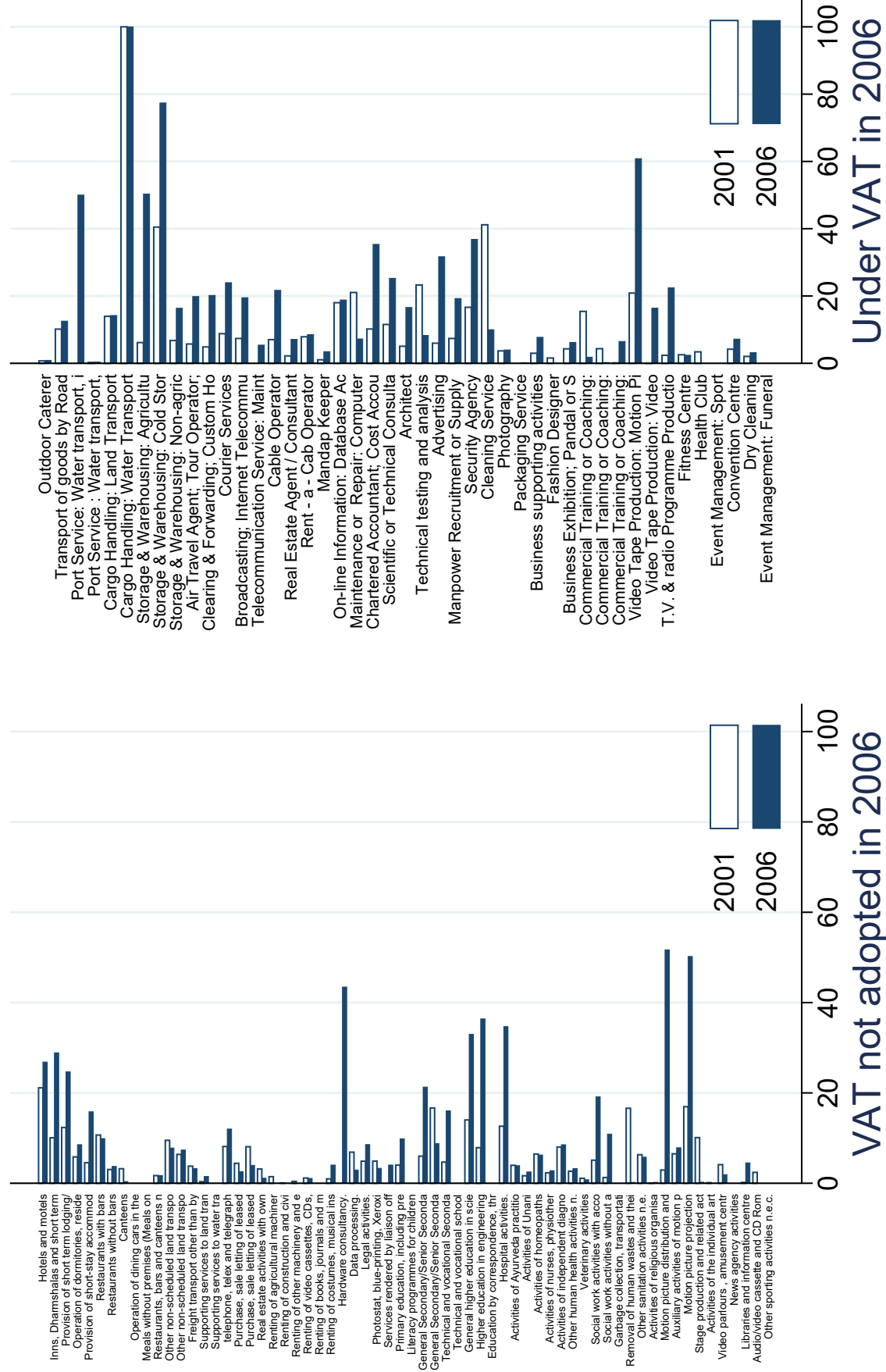


Figure 4: Tax registration rate of each activity before and after VAT adoption in 2003. The activities are divided based on the applicability of the service tax rule in 2006.

Table 3: Robustness tests – The estimated regressions is (30). The dependent variable is the share of formal firms in each activity, state, and year. In columns (1)-(4), instead of tax registration, formality is defined as registration under companies act. In columns (5)-(8), forward and backward linkages are computed using another input-output tables of Indian economy that is estimated for 2006. In the last four columns, to increase the number of services in the sample, the activities that service tax applied to them before 2001 are added.

	alternative formality definition				alternative linkages measure				adding activities under ST before 2001			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
VAT adoption	4.094*** (1.177)	-8.225** (3.313)	-2.532 (5.478)	-3.590 (5.409)	3.652*** (1.190)	-5.888** (2.798)	0.352 (4.759)	0.860 (4.719)	6.465*** (1.165)	-14.294*** (3.605)	-11.091** (4.335)	-14.317*** (4.619)
VAT adoption \times FL		5.164*** (1.641)		7.283*** (1.871)		3.326*** (1.186)		5.096*** (1.542)		7.489*** (1.756)		7.479*** (1.947)
VAT adoption \times BL			2.467 (3.061)	-4.990 (3.623)			0.546 (2.566)	-6.028* (3.304)			6.814*** (2.444)	0.024 (2.646)
Average no. workers	1.007*** (0.060)	1.006*** (0.059)	1.000*** (0.061)	1.011*** (0.060)	0.996*** (0.105)	0.997*** (0.107)	0.990*** (0.111)	1.000*** (0.107)	0.384*** (0.068)	0.371*** (0.068)	0.370*** (0.069)	0.371*** (0.067)
Number of firms	-0.025** (0.010)	-0.025** (0.011)	-0.024** (0.011)	-0.025** (0.011)	-0.025** (0.011)	-0.026** (0.011)	-0.026** (0.011)	-0.026** (0.011)	-0.018 (0.012)	-0.020 (0.013)	-0.019 (0.013)	-0.020 (0.013)
Observations	1119	1119	1119	1119	1119	1119	1119	1119	1631	1631	1631	1631
R-squared	0.404	0.440	0.434	0.442	0.376	0.400	0.397	0.402	0.307	0.340	0.334	0.340
Activity fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State \times year FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes